Testing the Constancy of Conditional Correlations in Multivariate GARCH-type Models (Extended Version with Appendix)

Anne Péguin-Feissolle, Bilel Sanhaji

Working Papers / Documents de travail

Testing the Constancy of Conditional Correlations in Multivariate GARCH-type Models (Extended Version with Appendix)

Anne Péguin-Feissolle
Bilel Sanhaji

WP 2015 - Nr 16

Testing the Constancy of Conditional Correlations in Multivariate GARCH-type Models (Extended Version with Appendix)

Anne Péguin-Feissolle
Aix Marseille University (Aix-Marseille School of Economics), CNRS & EHESS, Aix-Marseille, France (anne.peguin@univ-amu.fr)

Bilel Sanhaji
Aix Marseille University (Aix-Marseille School of Economics), CNRS & EHESS, Aix-Marseille, France (bilel.sanhaji@univ-amu.fr)

March 10, 2015

Abstract

We introduce two multivariate constant conditional correlation tests that require little knowledge of the functional relationship determining the conditional correlations. The first test is based on artificial neural networks and the second one is based on a Taylor expansion of each unknown conditional correlation. These new tests can be seen as general misspecification tests of a large set of multivariate GARCH-type models. We investigate the size and the power of these tests through Monte Carlo experiments. Moreover, we study their robustness to non-normality by simulating some models such as the $GARCH-t$ and $Beta-t-EGARCH$ models. We give some illustrative empirical examples based on financial data.

Key Words: Multivariate GARCH; Neural network; Taylor expansion

JEL classification: C22; C45; C58

1 Introduction

During the last decade, there has been a rapid increase in the amount of literature on theoretical and empirical derivatives of multivariate GARCH-type modeling (see, for example, extensive surveys of Bauwens, Laurent and Rombouts 2006 and Silvennoinen and Teräsvirta 2009a). This econometric modeling is an important issue, particularly in finance. Indeed, it is largely accepted that empirical models are more relevant in a multivariate than an univariate framework. By taking into account the possible dependence between individual equations, one can determine for example option pricing, asset pricing, hedging and risk management or portfolio selection, as noted in Bauwens, Laurent and Rombouts (2006). For instance, a crucial problem in financial studies is the correlation structure among different national stock returns because it determines the gains from international portfolio diversification. Although few tests are specific to multivariate models, it would be desirable to check *ex ante* the properties of the data especially since estimating multivariate GARCH (MGARCH) models is time-consuming. Moreover, it is worth noting that, if we don't test first the hypothesis of the constancy of conditional correlations, the estimation of some multivariate GARCH-type models could lead to estimation issues because some parameters could not be identified.

Most of the tests for the constancy of conditional correlations are based on the Lagrange Multiplier (LM) procedure. Some other tests are based on Likelihood Ratio procedure (see, among others, Longin and Solnik 1995 and Engle and Sheppard 2001). But, as noted by Tse (2000), computing the LM test statistics requires only the estimation of the constant correlation model and is thus generally computationally convenient. Tse (2000) introduces a LM test for the constant correlation hypothesis in a multivariate GARCH model; he extends the constant correlation model to one in which the correlations are allowed to be time-varying and he tests for the zero restrictions on the key parameters. Bera and Kim (2002) apply the White's (1982) and Orme (1990) information matrix (IM) test to the constant correlation bivariate GARCH model and establish the linkage between the IM and score tests. Silvennoinen and Teräsvirta (2015) derive an LM test for the constancy of correlations against the Smooth Transition Conditional Correlation (STCC) GARCH model in the multivariate case with an endogenous or exogenous transition variable. Berben and Jansen (2005) introduce a bivariate GARCH model with smoothed time-varying correlations based on logistic function, called the Smooth Transition Correlation GARCH (STCC-GARCH) model, and derive a new test for constant correlation, building on the LM test developed by Tse (2000). In contrast to Silvennoinen and Teräsvirta (2015), their model is bivariate and the variable controlling the transition between the extreme regimes is simply the time. Silvennoinen and Teräsvirta (2009b) present a test for the constancy of correlations against the DSTCC-GARCH (double smooth transition conditional correlation GARCH) model, i.e. by allowing the conditional correlations to vary according to two transition variables that can be stochastic or deterministic (for example, exogenous variables or lagged elements of endogenous variables).

In this paper, we introduce two LM tests for constancy of the conditional correlations.

The first test is based on artificial neural networks (ANN) or, more precisely, on single hidden layer perceptrons (later on, we use the generic term of artificial neural networks, following the terminology employed in the econometric literature). This ANN-based test relies on a statistical technique proposed by Lee, White and Granger (1993). The ANN framework has been already used for some tests: see, for example, Lee, White and Granger (1993) and Teräsvirta, Lin and Granger (1993) for linearity tests, Kamstra (1993), Caulet and Péguin-Feissolle (2000) and Péguin-Feissolle (1999) for conditional heteroscedasticity tests, Lebreton and Péguin-Feissolle (2007) for heteroscedasticity test and Péguin-Feissolle and Teräsvirta (1999) for causality tests.

The second test we introduce is based on a Taylor expansion of each unknown conditional correlation around a given point in a sample space; it is a way of linearizing the testing problem by approximating the true relationships by a Taylor series expansion. Thus, because of the linearization of the unknown relationship determining each conditional correlation, this test is not computationally more difficult to carry out than traditional tests. This kind of test has already been introduced in order to test the causality (Péguin-Feissolle and Teräsvirta 1999 and Péguin-Feissolle, Strikholm and Teräsvirta 2013), the heteroscedasticity (Lebreton and Péguin-Feissolle 2007) and the conditional heteroscedasticity (Caulet and Péguin-Feissolle 2000 and Péguin-Feissolle 1999).

Therefore, the two tests present four fundamental characteristics. First, they require little knowledge of the functional relationship determining the correlations. Secondly, they are easy to implement and perform well in our small-sample simulations; indeed, we show in the simulations that they are relevant in small samples and can be useful in investigating potential time-varying conditional correlation. Thirdly, they generalize well to high dimensions, i.e. with a high number of endogenous variables. Fourthly, given the properties of the ANN to be universal approximators and given that the Taylor expansion permits the linearization of an unknown relationship, a rejection of the null hypothesis of constancy of conditional correlations does not imply that the data have been generated from a model where the conditional correlations are specified as neural functions or specific functions; these tests can thus be seen as general misspecification tests of very different multivariate GARCH-type models. It is worthwhile to remark too that these tests could be used easily to test also the hypothesis of partially constant correlations.

Finite-sample properties of the two new tests are examined using Monte Carlo methods by comparing them to two alternative conditional correlation tests: the tests of Tse (2000) and Silvennoinen and Teräsvirta (2015). Using a variety of different specifications for the MGARCH model, we show that they perform well in our small-sample simulations; they approximate quite well unknown specifications of the conditional correlations, even in the case where the normality hypothesis is not verified. Empirical illustrations using real data point out that these tests can be useful to reject the constant correlation hypothesis in multivariate modelling of conditional heteroscedasticity.

The paper is organized as follows. In Section 2, we introduce the ANN-based constant correlation test and the Taylor-expansion based test. Section 3 reports the results of a simulation study; we investigate the size and the power by Monte Carlo experiments in small samples. In Section 4, we study the robustness of the tests to non-normality by simulating some models such as the $GARCH-t$ model of Bollerslev (1987), or the $Beta-t-EGARCH$ model of Harvey and Sucarrat (2014) and Harvey and Chakravarty (2008). Section 5 describes some illustrative empirical examples based on financial data. Section 6 contains some conclusions. Technical derivations of the test statistics presented in the paper can be found in the Appendix.

2 Testing the Constancy of Conditional Correlations

In order to fix the model that we consider and the notations, we present a general MGARCH model. Let $\{\mathbf{y}_t\}$ be a multivariate time series, where $\mathbf{y}_t$ is a $N \times 1$ vector. Consider the following model defined by, for $t = 1, ..., T$:

$$\begin{aligned} \mathbf{y}_t &= E[\mathbf{y}_t|\Omega_{t-1}] + \boldsymbol{\varepsilon}_t \\ Var\left[\mathbf{y}_t\left|\Omega_{t-1}\right.\right] &= \mathbf{H}_t \end{aligned} \tag{1}$$

where $E[\mathbf{y}_t|\Omega_{t-1}]$ and $Var\left[\mathbf{y}_t\left|\Omega_{t-1}\right.\right]$ are respectively the conditional expectation and the conditional covariance of $\mathbf{y}_t$ with respect to Ω_{t-1}, the sigma-field generated by all the information until time $t-1$. The process $\mathbf{y}_t$ is strictly stationary and ergodic. To simplify the discussion (but the test presented in this paper could be generalized easily to a large variety of time-varying structures of the conditional expectation of $\mathbf{y}_t$), we assume that the observations are of zero means: for $t = 1, ..., T$, $E_{t-1}[\mathbf{y}_t] = 0$ or

$$\mathbf{y}_t = \boldsymbol{\varepsilon}_t; \tag{2}$$

$\boldsymbol{\varepsilon}_t$ is defined by

$$\boldsymbol{\varepsilon}_t = \mathbf{H}_t^{1/2}\boldsymbol{\eta}_t \tag{3}$$

with $\boldsymbol{\eta}_t \sim iid\,(0, \mathbf{I}_N)$. We assume for simplicity that the conditional variances follow a GARCH(1,1) process:

$$h_{iit} = \zeta_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{ii,t-1} \qquad i = 1, ...N \tag{4}$$

where the standard positivity and covariance stationarity constraints are imposed, i.e. $\zeta_i > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$, and $\alpha_i + \beta_i < 1$ for $i = 1, ..., N$. Moreover, the conditional covariances are:

$$h_{ijt} = \rho_{ijt}\sqrt{h_{iit}h_{jjt}} \qquad 1 \leq i < j \leq N \tag{5}$$

where ρ_{ijt} is the conditional correlation. Let $\mathbf{P}_t = (\rho_{ijt})_{i,j=1,...,N}$ be the $N \times N$ conditional correlation matrix for the $\boldsymbol{\varepsilon}_t$; we can write

$$\mathbf{H}_t = \mathbf{S}_t\mathbf{P}_t\mathbf{S}_t \tag{6}$$

where $\mathbf{S}_t$ is the $N \times N$ matrix given by

$$\mathbf{S}_t = diag(\sqrt{h_{11t}}, ..., \sqrt{h_{NNt}}). \tag{7}$$

We assume that the conditional correlation matrix $\mathbf{P}_t$ is positive definite at each t; it guarantees the positive definiteness of $\mathbf{H}_t$. Moreover, we will assume the conditional normality of the $\boldsymbol{\varepsilon}_t$:

$$\boldsymbol{\varepsilon}_t|\Omega_{t-1} \sim N\left(0, \mathbf{H}_t\right) \tag{8}$$

which implies

$$\mathbf{z}_t|\Omega_{t-1} \sim N\left(0, \mathbf{P}_t\right), \tag{9}$$

with $\mathbf{z}_t = \mathbf{S}_t^{-1}\boldsymbol{\varepsilon}_t$.

2.1 The ANN-based test

The first test we present is an artificial neural network (ANN) based LM–type test. In order to build the test, we extend here the ideas of the tests of constant correlation hypothesis of Tse (2000), Silvennoinen and Teräsvirta (2014, 2009b) and Berben and Jansen (2005) to a model where the specification of the conditional correlations ρ_{ijt} is defined by a neural function. Artificial neural network models provide another way of dealing with situations where the functional form of a potential relationship between two variables is not assumed known in advance. As noted in Péguin-Feissolle (1999), the main theoretical support for the use of ANN models in the present context is the universal mapping theorem: it states that under mild regularity conditions, ANN models provide arbitrarily accurate approximations to nonlinear mappings (see Hornik, Stinchcombe and White 1989 and 1990, Hornik 1991, Stinchcombe and White 1989, Cybenko 1989, Carroll and Dickinson 1989, among others). Therefore, the test can be seen as a general misspecification test of a large set of multivariate GARCH-type models. Because the ANN are universal approximators, a rejection of the null hypothesis of constancy of conditional correlations does not imply that the data have been generated from a model where the conditional correlations are specified as neural functions.

We are going to specify the time-varying structure of the conditional correlations. We assume that the conditional correlations ρ_{ijt} are changing smoothly over time depending on a neural function as follows

$$\rho_{ijt} = \rho_{ij} + \sum_{m=1}^{p} \delta_{ijm}(1 + \exp\{-w'_{ijt}\gamma_{ijm}\})^{-1} \tag{10}$$

where $1 \leq i < j \leq N$, $p < \infty$, w_{ijt} and γ_{ijm} are $(2q+1) \times 1$ vectors, and w_{ijt} is given by

$$w_{ijt} = (1, \widetilde{w}'_{ijt})' = (1, \varepsilon_{i,t-1}, ..., \varepsilon_{i,t-q}, \varepsilon_{j,t-1}, ..., \varepsilon_{j,t-q})'. \tag{11}$$

The conditional correlations ρ_{ijt} should respect the conditions $|\rho_{ijt}| \leq 1$, $1 \leq i < j \leq N$ and $t = 1, ...T$, and the corresponding correlation matrix has to be positive semidefinite; but, because we test the null hypothesis of constant conditional correlation, i.e., $\rho_{ijt} = \rho_{ij}$, we simply assume that these restrictions are verified in a neighborhood of the null hypothesis (as noted by Tse 2000, p. 111).

For each couple of indices (i, j), the $2q+1$ input units of the network send signals, amplified or attenuated by weighting factors γ_{ijm}, to p hidden units (or hidden nodes) that sum up the signals and generate a squashing function; this function is assumed to be a logistic function. As previously mentioned, it is important to remark that neural functions of the form (10) may approximate arbitrary functions quite well, given a sufficiently large number p of hidden units and a suitable choice of the parameters δ_{ijm} and γ_{ijm}.

Following the definition (10) of the conditional correlations, the null hypothesis of constant conditional correlation, i.e., $\rho_{ijt} = \rho_{ij}$, can be formulated as:

$$H_{01} : \delta_{ijm} = 0,\ 1 \leq i < j \leq N,\ 1 \leq m \leq p. \tag{12}$$

Under H_{01}, the γ_{ijm}, for $i = 1,...,N-1$, $j = i+1,...,$ N and $m = 1,$ $...,p$, are not identified. For this reason, the conventional maximum likelihood theory for deriving the test procedure is not applicable, as noted in Lebreton and Péguin-Feissolle (2007). To solve this problem, we use the method given by Lee, White and Granger (1993), i.e. we draw the unidentified parameters randomly from a uniform distribution. Therefore, we generate the hidden unit weights, i.e. each γ_{ijm}, for $i = 1,...,N-1$, $j = i+1,...,$ N and $m = 1,$ $...,p$, randomly from the uniform $[-\mu, \mu]$ distribution. In our simulations, following Lee, White and Granger (1993), we choose $\mu = 2$.

The null hypothesis can be tested using the LM procedure. Let $\boldsymbol{\theta}$ the vector of all the parameters of the model, i.e. the 3×1 parameter vector for the conditional variances, $\boldsymbol{\omega}_i = (\zeta_i, \alpha_i, \beta_i)'$ for $i = 1,...,N$, the conditional correlation ρ_{ij} with $1 \leq i < j \leq N$ and the parameters in the neural function, δ_{ijm} for $1 \leq i < j \leq N$, $1 \leq m \leq p$ and $p < \infty$. Under standard regularity conditions, the test statistic that we call $NEURAL$ is given by:

$$\frac{1}{T}\left(\sum_{t=1}^{T}\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)' \Im(\boldsymbol{\theta})^{-1}\left(\sum_{t=1}^{T}\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right) \tag{13}$$

where $\Im(\boldsymbol{\theta})$ is replaced by the consistent estimator:

$$\Im(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T} E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right] \tag{14}$$

and the log-likelihood for the observation t, $l_t(\boldsymbol{\theta})$, is given by:

$$l_t(\boldsymbol{\theta}) = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln(|\mathbf{P}_t|) - \frac{1}{2}\sum_{i=1}^{N}\ln(h_{iit}) - \frac{1}{2}\mathbf{z}_t'\mathbf{P}_t^{-1}\mathbf{z}_t. \tag{15}$$

In practice (see Tse 2000), $\Im(\boldsymbol{\theta})$ may be replaced using the negative of the Hessian matrix:

$$\Im(\boldsymbol{\theta}) = -\frac{1}{T}\sum_{t=1}^{T} E_{t-1}\left[\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right]. \tag{16}$$

The statistic $NEURAL$ has an asymptotic χ^2 distribution under the null with $p\frac{N(N-1)}{2}$ degrees of freedom. Appendix presents the details of the technical derivations of the test statistic.

As Lee, White and Granger (1993) and Péguin-Feissolle and Teräsvirta (1999) pointed out, the elements of some matrices used to build the ANN-based test statistic can lead to collinearity problems when the number of hidden units p is large. Therefore, in our case, the $p \times 1$ vectors $\mathbf{g}_{ijt}$ defined as, for $1 \leq i < j \leq N$, $p < \infty$:

$$\mathbf{g}_{ijt} = \begin{pmatrix} g_{ijt,1} \\ \vdots \\ g_{ijt,p} \end{pmatrix} = \begin{pmatrix} (1+\exp\{-w_{ijt}'\gamma_{ij1}\})^{-1} \\ \vdots \\ (1+\exp\{-w_{ijt}'\gamma_{ijp}\})^{-1} \end{pmatrix} \tag{17}$$

tend to be collinear among themselves especially when p is large. The conditional correlations ρ_{ijt} given by (10) can be written as follows

$$\rho_{ijt} = \rho_{ij} + \sum_{m=1}^{p} \delta_{ijm} g_{ijt,m}.$$

Let us define $\mathbf{G}_{ij}$ the $p \times T$ matrix given by:

$$\mathbf{G}_{ij} = (\mathbf{g}_{ij1}, ..., \mathbf{g}_{ijT}).$$

Thus we conduct the test using the main principal components of each $\mathbf{G}_{ij}$ matrix. Only the largest principal components that together explain at least 90% of the variation in this matrix are used; the number of principal components are determined automatically for each i and j (see Péguin-Feissolle and Teräsvirta 1999, Lebreton and Péguin-Feissolle 2007 and Péguin-Feissolle, Strikholm and Teräsvirta 2013; see also Castle and Hendry 2010 for discussions on using principal components to solve the collinearity problem). More precisely, instead of $\mathbf{G}_{ij}$, we build the $p_{ij}^* \times T$ matrix $\mathbf{G}_{ij}^*$ of the p_{ij}^* principal components chosen according to the preceding rule, i.e.

$$\mathbf{G}_{ij}^* = \left(\mathbf{g}_{ij1}^*, ..., \mathbf{g}_{ijT}^*\right).$$

The null hypothesis is now that the parameters associated to the main principal components are equal to zero in the following specification of the conditional correlations:

$$\rho_{ijt} = \rho_{ij} + \sum_{m=1}^{p_{ij}^*} \delta_{ijm} g_{ijt,m}^*.$$

The statistic $NEURAL$ will have in this case an asymptotic χ^2 distribution under the null with $\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} p_{ij}^{*}$ degrees of freedom.

2.2 The Taylor expansion-based test

We assume that the functional form f_{ij} determining ρ_{ijt}, for $i, j = 1, ..., N$ and $t = 1, ..., T$, is unknown and is adequately represented by the following equation:

$$\rho_{ijt} = f_{ij}\left(\widetilde{w}_{ijt}, \boldsymbol{\theta}_{ij}^{*}\right) \tag{18}$$

where, as given in (11), for each couple of indices (i, j), $\widetilde{w}_{ijt}$ is a $2q \times 1$ vector given by

$$\widetilde{w}_{ijt} = (\varepsilon_{i,t-1}, ..., \varepsilon_{i,t-q}, \varepsilon_{j,t-1}, ..., \varepsilon_{j,t-q})' \tag{19}$$

and $\boldsymbol{\theta}_{ij}^{*}$ is a $r_{\boldsymbol{\theta}_{ij}^{*}} \times 1$ unknown parameter vector. The conditional correlations ρ_{ijt} should respect the conditions $|\rho_{ijt}| \leq 1$, $1 \leq i < j \leq N$ and $t = 1, ...T$, and the corresponding correlation matrix has to be positive semidefinite; but, like in the ANN-based test, we will simply assume that these restrictions are verified in a neighborhood of the null hypothesis.

The test is based on a finite-order Taylor expansion. Following Péguin-Feissolle and Teräsvirta (1999) and Péguin-Feissolle, Strikholm and Teräsvirta (2013), we assume that all the functions f_{ij} have a convergent Taylor expansion at any arbitrary point of the sample space for every $\boldsymbol{\theta}_{ij}^{*} \in \Theta_{ij}$ (the parameter spaces) in order to ensure that, when the order of the Taylor expansion increases, the remainder of the Taylor expansion converges to zero, for each couple (i, j).

In order to linearize now f_{ij} in (18), we expand the function into a $k$thorder Taylor series around an arbitrary fixed point in the sample space. After approximating f_{ij}, merging terms and reparametrizing, we obtain:

$$\begin{aligned}\rho_{ijt} = \;& \rho_{ij} + \sum_{m=1}^{q} \lambda_m \varepsilon_{i,t-m} + \sum_{m=1}^{q} \phi_m \varepsilon_{j,t-m} + \sum_{m_1=1}^{q} \sum_{m_2=m_1}^{q} \lambda_{m_1 m_2} \varepsilon_{i,t-m_1} \varepsilon_{i,t-m_2} \\ & + \sum_{m_1=1}^{q} \sum_{m_2=1}^{q} \psi_{m_1 m_2} \varepsilon_{i,t-m_1} \varepsilon_{j,t-m_2} + \sum_{m_1=1}^{q} \sum_{m_2=m_1}^{q} \phi_{m_1 m_2} \varepsilon_{j,t-m_1} \varepsilon_{j,t-m_2} \\ & + ... + \sum_{m_1=1}^{q} \sum_{m_2=m_1}^{q} ... \sum_{m_k=m_{k-1}}^{q} \lambda_{m_1 ... m_k} \varepsilon_{i,t-m_1} ... \varepsilon_{i,t-m_k} \\ & + ... + \sum_{m_1=1}^{q} \sum_{m_2=m_1}^{q} ... \sum_{m_k=m_{k-1}}^{q} \phi_{m_1 ... m_k} \varepsilon_{j,t-m_1} ... \varepsilon_{j,t-m_k} + R_t^{(k)}\end{aligned} \tag{20}$$

where $R_t^{(k)}$ is the remainder; in the expansion (20) we find all possible combinations of lagged values of $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$.

The assumption that the conditional correlation ρ_{ijt} is constant means that all terms involving functions of elements of lagged values of $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$ in (20) must have zero coefficients, i.e. all the parameters except ρ_{ij} are equal to zero. Therefore, the null hypothesis of interest is:

$$H_{02} : \begin{cases} \lambda_m = 0, \; m = 1, ..., q \\ \phi_m = 0, \; m = 1, ..., q \\ \lambda_{m_1 m_2} = 0, \; m_1 = 1, ..., q, \; m_2 = m_1, ..., q \\ \psi_{m_1 m_2} = 0, \; m_1 = 1, ..., q, \; m_2 = 1, ..., q \\ \phi_{m_1 m_2} = 0, \; m_1 = 1, ..., q, \; m_2 = m_1, ..., q \\ \vdots \\ \phi_{m_1 ... m_k} = 0, \; m_1 = 1, ..., q, \; m_2 = m_1, ..., q, ..., \; m_k = m_{k-1}, ..., q. \end{cases}$$

The number of parameters to be tested under the null hypothesis is:

$$N^* = 2\left[\sum_{r=1}^{k}\left(\begin{array}{c} 2q+r-1 \\ r \end{array}\right) - \sum_{r=1}^{k}\left(\begin{array}{c} q+r-1 \\ r \end{array}\right)\right]. \tag{21}$$

The null hypothesis can be tested based on the LM procedure as before in the case of the $NEURAL$ statistic. Under standard regularity conditions, the test statistic $TAYLOR$ is given by formula (13) where $\boldsymbol{\theta}$ is the vector of all the parameters of the model, i.e. the parameters for the conditional variances $\boldsymbol{\omega}_i = (\zeta_i, \alpha_i, \beta_i)'$ for $i = 1$, ..., N, the conditional correlations ρ_{ij} with $1 \leq i < j \leq N$ and the $N^* \times 1$ parameters that are equal to zero under the null hypothesis. When the remainder $R_t^{(k)} \equiv 0$, the statistic $TAYLOR$ has an asymptotic χ^2 distribution with N^* degrees of freedom under the null hypothesis. Appendix presents the details of the technical derivations of the test statistic.

Following Péguin-Feissolle, Strikholm and Teräsvirta (2013), there are two practical difficulties when the order of the Taylor expansion k is increasing, first the regressors tend to be highly collinear, and second the dimension of the null hypothesis may become rather large because the number of these regressors increases rapidly with k. More precisely, the conditional correlations ρ_{ijt} given in (20) can be written as

$$\rho_{ijt} = \rho_{ij} + \sum_{m=1}^{N^*} \delta_{ijm} d_{ijt,m} + R_t^{(k)}$$

where the $N^* \times 1$ vectors $\mathbf{d}_{ijt}$ correspond to all terms involving functions of elements of lagged values of $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$. We replace the matrix composed by the $N^* \times 1$ vectors $\mathbf{d}_{ijt}$ by its largest principal components: like in the case of the ANN-based test, only the largest principal components that together explain at least 90% of the variation in the matrix are used, the number of principal components being determined automatically for each i and j. Instead of the $N^* \times T$ matrix D_{ij} composed by the $\mathbf{d}_{ijt}$ vectors:

$$\mathbf{D}_{ij} = (\mathbf{d}_{ij1}, ..., \mathbf{d}_{ijT})$$

for $1 \leq i < j \leq N$, we build the $N_{ij}^* \times T$ matrix $\mathbf{D}_{ij}^*$ of the N_{ij}^* chosen principal components, i.e.

$$\mathbf{D}_{ij}^* = \left(\mathbf{d}_{ij1}^*, ..., \mathbf{d}_{ijT}^*\right).$$

The null hypothesis will be that the parameters associated to the main principal components are equal to zero in the following relationship determining ρ_{ijt}:

$$\rho_{ijt} = \rho_{ij} + \sum_{m=1}^{N_{ij}^*} \delta_{ijm} d_{ijt,m}^*.$$

The statistic $TAYLOR$ will follow an asymptotic χ^2 distribution under the null with $\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} N_{ij}^*$ degrees of freedom.

3 Monte Carlo experiments

In this section, we investigate the small-sample performances of the ANN-based test and the Taylor expansion-based test for the constancy of conditional correlations in multivariate GARCH-type models. By using Monte-Carlo experiments, we compare both the sizes and the powers of these tests to two alternative conditional correlation tests: the tests of Tse (2000) and Silvennoinen and Teräsvirta (2015).

3.1 Simulation design

The data generating process (DGP) is based on the general multivariate GARCH model introduced before where the $N \times 1$ vector of residuals is given by

$$\boldsymbol{\varepsilon}_t = \mathbf{H}_t^{1/2} \boldsymbol{\eta}_t$$

with $\boldsymbol{\eta}_t \sim nid\,(0, \mathbf{I}_N)$, $t = 1, \ldots, T$. In other words, we generate several MGARCH-type models generally used in financial time series analysis, based on the normality assumption; we will relax the normality assumption in Section 4. For each simulation, $\boldsymbol{\eta}_t$ thus follows a multivariate Gaussian distribution with zero mean vector and covariance matrix $\mathbf{I}_N$. For each model, the DGP is used to compute the empirical sizes and powers in order to compare the performances of both the tests we propose (ANN-based test and Taylor expansion-based test) to two well-known tests (Tse (2000) test and Silvennoinen and Teräsvirta (2015) test).

For all the Monte Carlo simulations, the sample sizes are $T = 1000$, 1500 and 2500 and the number of endogenous variables is $N = 2$ (except in Table 6 where $N = 3$). Moreover, we remove the first observations in order to eliminate initialization effects. The number of replications is $S = 2000$ (except in Tables 1 and 2). Concerning the artificial neural network based tests, following Lee, White and Granger (1993), the number of hidden units is $p = 20$; moreover, we generate the hidden unit weights, i.e. the different

γ_{ijm} in (10), for $1 \leq i < j \leq N$ and $1 \leq m \leq p$, randomly from the uniform $[-\mu, \mu]$ distribution. Following the same authors, we choose $\mu = 2$. For a fixed number of lags of residuals in the case of the artificial neural network based test and for a fixed number of lags of residuals and a fixed order of the Taylor expansion in the case of the Taylor expansion-based test, only the largest principal components that together explain at least 90% of the variation in the corresponding matrices are used, the number of principal components being determined automatically.

[**Insert Table 1 here**]

[**Insert Table 2 here**]

The performances of the two new tests are compared to two well-known tests:

Tse (2000) test: TSE denotes the LM statistic developed by Tse (2000) given by (we report the paper's notations, see Tse 2000 for further details):

$$TSE = \mathbf{s}'(\mathbf{S}'\mathbf{S})^{-1}\mathbf{s} \tag{22}$$

where $\mathbf{s} = \mathbf{S}'\mathbf{l}$, $\mathbf{l}$ is the $T \times 1$ column vector of ones and $\mathbf{S}$ is the $T \times N$ matrix of which the rows are the partial derivatives $\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}$ for $t = 1, .., T$; $l_t(\boldsymbol{\theta})$ is given by (15). $\boldsymbol{\theta}$ is evaluated at $\widehat{\boldsymbol{\theta}}$ which is the maximum likelihood estimator of $\boldsymbol{\theta}$ under the null hypothesis.

Silvennoinen and Teräsvirta (2015) test: these authors assume that the conditional correlation matrix $\mathbf{P}_t$ is given by

$$\mathbf{P}_t = (1 - \mathbf{G}_t)\mathbf{P}_1 + \mathbf{G}_t\mathbf{P}_2$$

where $\mathbf{P}_1$ and $\mathbf{P}_2$ are positive definite correlation matrices; $\mathbf{G}_t$ is a logistic transition function:

$$\mathbf{G}_t = \left(1 + e^{-\gamma(\mathbf{s}_t - c)}\right)^{-1}$$

where γ is the transition parameter ($\gamma > 0$), s_t is the transition variable and c is the location parameter. Using a first-order Taylor approximation, $\mathbf{P}_t$ can be approximated as $\mathbf{P}_t^* = \mathbf{P}_1^* - \mathbf{s}_t\mathbf{P}_2^*$ where $\mathbf{P}_1^*$ and $\mathbf{P}_2^*$ are depending on γ, $\mathbf{P}_1$ and $\mathbf{P}_2$. The null hypothesis can be written $vecl(\mathbf{P}_2^*) = 0$, where the $vecl$ operator stacks the columns of the strict lower diagonal of the square argument matrix (i.e. by excluding the diagonal elements). The LM statistic, $STCC$, is given by (13) where $l_t(\boldsymbol{\theta})$ is given by (15). $\boldsymbol{\theta}$ is evaluated at $\hat{\boldsymbol{\theta}}$ which is the maximum likelihood estimator of $\boldsymbol{\theta}$ under the null hypothesis (see Silvennoinen and Teräsvirta 2014 for more details).

3.2 Size simulations

We present here the results concerning the size of the different tests, i.e. in the case where the data are generated under the null hypothesis using a constant conditional correlation model. Following Silvennoinen and Teräsvirta (2015), the transition variable for the $STCC$ test will be generated in the simulations from an exogenous $GARCH(1,1)$ process such that $\mathbf{s}_t = \mathbf{h}_t^{1/2}\mathbf{z}_t$ where $\mathbf{z}_t \sim N(0,1)$ and $\mathbf{h}_t = 0.02 + 0.03\mathbf{s}_{t-1}^2 + 0.94\mathbf{h}_{t-1}$.

We thus consider the following models. The first one is an extended CCC-GARCH model (ECCC-GARCH) defined by Jeantheau (1998). It is a generalization of the multivariate GARCH model with constant conditional correlations (CCC-GARCH) proposed by Bollerslev (1990) and Baillie and Bollerslev (1990); the individual conditional variance equations depend on the past squared returns and conditional variances of all series. Therefore, the elements of the conditional variance matrix are characterized for $N = 2$ by:

$$\begin{pmatrix} h_{11t} \\ h_{22t} \end{pmatrix} = \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} \varepsilon^2_{1,t-1} \\ \varepsilon^2_{2,t-1} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} h_{11,t-1} \\ h_{22,t-1} \end{pmatrix} \tag{23}$$

and

$$h_{12t} = \rho_{12}\sqrt{h_{11t}h_{22t}} \tag{24}$$

where ρ_{12} is the constant conditional correlation. This model will be denoted by $ECCC$. In the second model, called CCC, the off-diagonal parameters in (23) are equal to zero i.e. $\alpha_{12} = \alpha_{21} = 0$ and $\beta_{12} = \beta_{21} = 0$. The third model, called $GARCH$, will be composed of two univariate independent GARCH(1,1) models, i.e. $\alpha_{12} = \alpha_{21} = 0$, $\beta_{12} = \beta_{21} = 0$ and $\rho_{12} = 0$ in (23) and (24). We also design an asymmetric model where (23) is completed by leverage coefficients such that

$$\begin{pmatrix} h_{11t} \\ h_{22t} \end{pmatrix} = \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} \varepsilon^2_{1,t-1} \\ \varepsilon^2_{2,t-1} \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} \nu^2_{1,t-1} \\ \nu^2_{2,t-1} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} h_{11,t-1} \\ h_{22,t-1} \end{pmatrix} \tag{25}$$

where $\nu_{i,t} = I(\varepsilon_{i,t} < 0)\varepsilon_{i,t}$ for $i = 1, \ldots, N$. The indicator function $I()$ takes value one if the argument is true and zero otherwise. Thus, the leverage coefficients are applied to negative innovations, giving additional weights to negative changes. The fourth model denoted GJR is composed by two univariate independent GJR-GARCH models, i.e. $\alpha_{12} = \alpha_{21} = 0$, $\gamma_{12} = \gamma_{21} = 0$, $\beta_{12} = \beta_{21} = 0$ and $\rho_{12} = 0$ in (24) and (25).

To determine the optimal order of the Taylor expansion (k) and the optimal number of lags of residuals (q) used in both the new tests, we generate $S = 1000$ replications of the ECCC and the bivariate GARCH models. Tables 1 and 2 summarize the 1%, 5% and the 10% rejection frequencies for various q lags of residuals and k orders of Taylor expansion for these models. The selection criterion is based on the size distortions of the both tests from the nominal sizes for different sample sizes. Regarding the simulation results, among other possibilities, the choice of $q = 3$ for $NEURAL$ seems to be the best choice. The test statistic $NEURAL$ is therefore simulated for a number of lags q in (11) equal to 3. Moreover, the order of the Taylor expansion $k = 3$ accompanied by a number of lags in (19) $q = 2$ is likely to be a good choice, leading to $N^* = 50$ before the main component analysis. Therefore, we summarize the different parameters for the two tests that will be used in the rest of the article as follows:

- for the test statistic $NEURAL$: the number of hidden units is $p = 20$ and the number of lags of residuals is $q = 3$,

- for the test statistic $TAYLOR$: the order of the Taylor expansion is $k = 3$ and the number of lags of residuals is $q = 2$.

Table 3 shows the empirical sizes of the different tests, i.e. the rejection probabilities under the null hypothesis of constant conditional correlations, assuming nominal sizes of 1%, 5% and 10% and different sample sizes ($T =$ 1000, 1500 and 2500). Moreover, we consider models with different constant correlations: high with $\rho_{12} = 0.60$ for the ECCC-GARCH, medium with $\rho_{12} =$ 0.30 for the CCC model, and of course $\rho_{12} = 0$ for the $GARCH$ and GJR models. The DGP models correspond to different persistences, for example $\alpha_{11} + \beta_{11} = \alpha_{22} + \beta_{22} = 0.99$ for the CCC-GARCH model, and $\alpha_{11} + \beta_{11} = 0.95$ and $\alpha_{22} + \beta_{22} = 0.90$ in the bivariate GARCH model.

[**Insert Table 3 here**]

For all the tests, we can observe that sizes converge towards the nominal sizes when the number of observations T increases, for most of the tests. More precise conclusions can be given.

In the $ECCC$ case, for the different sample sizes, the TSE test statistic over-rejects the null hypothesis while a multivariate constant conditional correlation model is generated. The $NEURAL$, the $TAYLOR$ and the $STCC$ test statistics show similar performances and clearly outperform the TSE test.

In the CCC case, the sizes seem close to the nominal sizes for all the test statistics when the sample sizes reach $T = 2500$.

The bivariate $GARCH$ case permits to study the possibility of no correlation at all, i.e. the nullity of the conditional covariances. The rejection frequencies show clear signs of over-rejections for all the sample sizes for the TSE test statistic; the $NEURAL$, $TAYLOR$ and $STCC$ test statistics clearly outperform the TSE test statistics, which shed light on their good size properties.

The bivariate GJR case allows positive and negative shocks to have an asymmetric effect in the conditional variances. The $NEURAL$, the $TAYLOR$ and the $STCC$ tests show good properties for $T = 2500$, while the TSE test tends to over-reject.

3.3 Power simulations

To illustrate the behavior of all tests under the alternative hypothesis, we generate different time-varying conditional correlations multivariate GARCH-type models, chosen to represent a variety of situations. It is important to note that the performance of the $STCC$ test depends on the choice of the transition variable, but following Silvennoinen and Teräsvirta (2015), we define the transition variable for the $STCC$ test as a linear combination of lags of squared returns: $\mathbf{s}_t = (0.2, 0.2, 0.2, 0.2, 0.2) \times \left(\overline{\boldsymbol{\varepsilon}}^2_{t-1}, \overline{\boldsymbol{\varepsilon}}^2_{t-2}, \overline{\boldsymbol{\varepsilon}}^2_{t-3}, \overline{\boldsymbol{\varepsilon}}^2_{t-4}, \overline{\boldsymbol{\varepsilon}}^2_{t-5}\right)'$ where $\overline{\boldsymbol{\varepsilon}}^2_t$ is the mean of $\boldsymbol{\varepsilon}_t$ over N of its squared elementwise. We consider the following models.

The $BEKK(1,1,1)$ model (Engle and Kroner 1995) is defined by:

$$\mathbf{H}_t = \mathbf{Z}'\mathbf{Z} + \mathbf{A}'\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}'_{t-1}\mathbf{A} + \mathbf{B}'\mathbf{H}_{t-1}\mathbf{B} \tag{26}$$

where $\mathbf{Z}$, $\mathbf{A}$ and $\mathbf{B}$ are $N \times N$ matrices and $\mathbf{Z}$ is upper triangular. With $N = 2$, we have:

$$\mathbf{Z} = \begin{pmatrix} \zeta_{11} & \zeta_{12} \\ 0 & \zeta_{22} \end{pmatrix}, \mathbf{A} = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \text{ and } \mathbf{B} = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \tag{27}$$

Note that the conditional covariance matrices $\mathbf{H}_t$ are positive definite by construction. This model will be denoted by $BEKK$.

An asymmetric scalar $BEKK$ model is also tested (Ding and Engle 2001). This model is a particular case of the $BEKK$ model (27) and will be denoted by $ASBEKK$:

$$\mathbf{H}_t = \mathbf{ZZ}' + \alpha\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}'_{t-1} + \gamma\boldsymbol{\nu}_{t-1}\boldsymbol{\nu}'_{t-1} + \beta\mathbf{H}_{t-1}, \tag{28}$$

where α, γ and β are scalars and $\boldsymbol{\nu}_t = \mathbf{I}(\boldsymbol{\varepsilon}_t < 0) \odot \boldsymbol{\varepsilon}_t$. $\mathbf{I}()$ is the $N \times 1$ indicator function and $\odot$ is the Hadamard product.

Another model is the DCC-GARCH model of Engle (2002) (denoted by DCC): $\mathbf{H}_t$ is defined as in (6), i.e. $\mathbf{H}_t = \mathbf{S}_t\mathbf{P}_t\mathbf{S}_t$ with $\mathbf{P}_t$ the $N \times N$ conditional correlation matrix and $\mathbf{S}_t$ is the $N \times N$ matrix given by $\mathbf{S}_t = diag(\sqrt{h_{11t}}, ..., \sqrt{h_{NNt}})$, with

$$\mathbf{P}_t = diag\left(q_{11,t}^{-1/2}, ..., q_{NN,t}^{-1/2}\right)\mathbf{Q}_t diag\left(q_{11,t}^{-1/2}, ..., q_{NN,t}^{-1/2}\right) \tag{29}$$

where $\mathbf{Q}_t = (q_{ij,t})$ is a symmetric positive definite $N \times N$ matrix given by

$$\mathbf{Q}_t = (1 - \alpha - \beta)\,\bar{\mathbf{Q}} + \alpha\mathbf{z}_{t-1}\mathbf{z}'_{t-1} + \beta\mathbf{Q}_{t-1}; \tag{30}$$

$\bar{\mathbf{Q}}$ is the $N \times N$ constant matrix (see Aielli 2013 for more details on $\bar{\mathbf{Q}}$), $\mathbf{z}_t = \mathbf{S}_t^{-1}\boldsymbol{\varepsilon}_t$ and α and β are non-negative scalar parameters satisfying $\alpha + \beta < 1$. Moreover, h_{iit} is a simple $GARCH(1,1)$ model for $i = 1, .., N$.

The last model is the EDCC-GARCH model (denoted by $EDCC$); it is the extended specification of DCC-GARCH model, allowing for volatility spillovers, i.e. the individual conditional variance equations depend on the past squared returns and variances of all series (see formula (23)). In the case

$N = 2$, we have for $i, j = 1, 2$:

$$h_{iit} = \zeta_{ii} + \alpha_{ii}\varepsilon_{i,t-1}^2 + \alpha_{ij}\varepsilon_{j,t-1}^2 + \beta_{ii}h_{ii,t-1} + \beta_{ij}h_{jj,t-1}. \tag{31}$$

We take different data generation processes for three of these models ($BEKK_1$ and $BEKK_2$, DCC_1 and DCC_2, $EDCC_1$ and $EDCC_2$), characterized by different generation parameters and different variabilities of conditional correlation coefficients. Table 4 gives the parameter values used to generate the different models; following Tse (2000), we give in the second part of the Table 4 the variability of the conditional correlation coefficients by calculating the range (maximum - minimum) of these coefficients in each simulated sample. The variability is high with DCC_2 and $EDCC_1$, moderate with $BEKK_2$ and DCC_1 and low with $BEKK_1$, $ASBEKK$ and $EDCC_2$. Table 5 summarizes the empirical powers of the ANN-based test, the Taylor expansion-based test, the TSE test and the $STCC$ test. In other words, it shows the rejection probabilities of the null hypothesis of constant conditional correlations assuming nominal sizes of 1%, 5% and 10% and different sample sizes ($T = 1000$, 1500 and 2500). For all the tests, we can observe in Table 5 that the rejection probabilities of the null hypothesis increase with the number of observations T.

[**Insert Table 4 here**]

[**Insert Table 5 here**]

For the three BEKK-GARCH specifications, $BEKK_1$, $BEKK_2$ and $ASBEKK$, the tests reject the null hypothesis of constancy of conditional correlation; nevertheless, even if most of them present good properties with a rejection probability sometimes equal to 1, the TSE test seems to have less good results, especially for $T = 1000$.

The DCC-GARCH models, DCC_1 and DCC_2, are characterized by parameters implying very different ranges of variability of the conditional correlation coefficients: the intervals are very large in the case of the DCC_2 model. Nevertheless, the $TAYLOR$ test presents the best performance in both DGP models followed by the $NEURAL$ test in the first one and the TSE test in the second one. On the other hand, the tests performing worse than the other tests are the TSE and $STCC$ tests in the DCC_1 model and the $NEURAL$ and $STCC$ tests in the DCC_2 model.

Concerning the extended DCC-GARCH models, $EDCC_1$ and $EDCC_2$, the tests showing the best performances are the $TAYLOR$ and TSE tests in the $EDCC_1$ case and the $NEURAL$ test in the $EDCC_2$ case. In the first case, the $NEURAL$, $TAYLOR$ and TSE tests outperform the $STCC$ test. In the second DGP model, the performances are altogether poor for each test.

Whatever the variability is, at least one of the new tests seems to present good quality in order to conclude to the non constancy of the conditional correlations. Concerning the advantages of the Neural test with respect to the Taylor test, the Neural test seems to have better properties when the variability of the conditional correlation coefficients is low ($EDCC_2$); it is important to note that in this last case, the Neural test is the only test that can detect the non constancy of conditional correlations. At the opposite, the Taylor test seems better in the case where the variability of the conditional correlation coefficients is high (DCC_2 and $EDCC_1$). Therefore, we can conclude that even if one or the other of the new tests is not the best each time, it is important to apply jointly both of them because at least one of them can detect the time-varying conditional correlations. The idea is that, if at least one of the two tests rejects the null hypothesis of constant conditional correlations, we can try a more complex model with time-varying conditional correlations.

We also generalize the $NEURAL$, $TAYLOR$, TSE and $STCC$ tests to the case of three endogenous variables ($N = 3$). Table 6 summarizes the results of the small sample sizes ($GARCH_2$) and powers ($BEKK_3$). All the tests, except the TSE test for the small sample powers, show good results because the sizes are close to the nominal sizes (see $GARCH_2$) and the powers are high (see $BEKK_3$).

[**Insert Table 6 here**]

All the multivariate GARCH-type models have thus been tested satisfactorily using the ANN-based test and the Taylor expansion-based test compared to TSE and $STCC$ tests. This finding suggests that whatever the GARCH-type model, both tests developed in this paper show relevant properties to identify time-varying conditional correlations. To summarize our results, the ANN-based and Taylor expansion-based tests permit to generalize the acceptance or rejection of the hypothesis of constant correlations in a general multivariate GARCH framework.

4 Robustness to non-normality

The tests developed in this paper have been studied in the above section with a DGP following a Gaussian distribution. In order to investigate the robustness of the different tests to non-normality, we simulate in this section some models with a non-normal conditional distribution: the multivariate versions of $GARCH-t$ model of Bollerslev (1987) and $Beta-t-EGARCH$ model of Harvey and Sucarrat (2014) and Harvey and Chakravarty (2008). In all the Monte Carlo simulations, the number of replications is $S = 2000$, the sample sizes are $T = 1000$, 1500 and 2500 and the number of endogenous variables is $N = 2$; we consider the nominal sizes of 1%, 5% and 10%.

4.1 Size simulations

We first study the empirical sizes of the different tests when the following two models are generated with null correlations. The first model corresponds to two independent univariate Student t–GARCH models; they are characterized by $\alpha_{12} = \alpha_{21} = 0$, $\beta_{12} = \beta_{21} = 0$ and $\rho_{12} = 0$ in (23), and thus $h_{12t} = 0$. The second model consists of two independent univariate $Beta-t-EGARCH$ with and without a leverage effect. We use the same notations as Harvey and Sucarrat (2014) who developed a more general framework than Harvey and Chakravarty (2008). Each $Beta-t-EGARCH$ model without a leverage effect is defined as follows, for $i = 1, 2$:

$$y_{it} = \mu_i + \varepsilon_{it} \exp\left(\lambda_{i,t|t-1}\right), \; t = 1, \ldots, T, \tag{32}$$

where ε_{it} has a t_{ν_i}–distribution and is serially independent and ν_i, the number of degrees of freedom, is positive. Let us define the conditional score

$$u_{it} = \frac{(\nu_i + 1)(y_{it} - \mu_i)^2}{\nu_i \exp(2\lambda_{i,t|t-1}) + (y_{it} - \mu_i)^2} - 1, \; -1 \leq u_{it} \leq \nu_i, \; \nu_i > 0. \tag{33}$$

We consider the first order model given by

$$\lambda_{i,t|t-1} = \delta_i + \phi_i \lambda_{i,t-1|t-2} + \kappa_i u_{i,t-1}; \tag{34}$$

with the stationarity condition $|\phi_i| < 1$. The first order $Beta-t-EGARCH$ with a leverage effect is defined in Harvey and Sucarrat (2014) as follows

$$\lambda_{i,t|t-1} = \delta_i + \phi_i \lambda_{i,t-1|t-2} + \kappa_i u_{i,t-1} + \kappa_i^* sign(-(y_{i,t-1} - \mu_i))(u_{i,t-1} + 1). \tag{35}$$

As noted in Harvey and Sucarrat (2014), if an increase in the absolute values of a standardized observation leads to an increase in volatility the restriction $\kappa_i \geq \kappa_i^* \geq 0$ may be imposed (see Harvey and Chakravarty 2008 and Harvey and Sucarrat 2014 for a complete definition of these models and their properties).

The generation parameters are given in Table 7. Table 8 shows the small sample sizes of $t-GARCH$ and $Beta-t-EGARCH$. When the DGP follows a Student distribution, the results of $t-GARCH$ model highlight the relative robustness to non-normality of the $TAYLOR$, $NEURAL$ and $STCC$ tests compared to the TSE test, even for small numbers of observations. When the observations are simulated with a first order $Beta-t-EGARCH$ model with or without leverage effects, the TSE test over-rejects the null hypothesis of constant conditional correlations. The $NEURAL$, $TAYLOR$ and $STCC$ tests are very close to the nominal sizes and show again better results. Generally, the $TAYLOR$, $NEURAL$ and $STCC$ tests outperform thus the TSE test.

[**Insert Table 7 here**]

[**Insert Table 8 here**]

4.2 Power simulations

We study the empirical powers in small samples with the same models studied for the sizes except that the conditional correlations are no longer constant. The conditional expectation model can be written, for $t = 1, \ldots, T$, and $i = 1$, ..., N,

$$y_{it} = \sqrt{h_{iit}}\varepsilon_{it} \tag{36}$$

where, for the $GARCH(1,1)$ models,

$$h_{iit} = \zeta_i + \alpha_i \varepsilon^2_{i,t-1} + \beta_i h_{ii,t-1} \tag{37}$$

and for the $Beta - t - EGARCH$ models,

$$h_{iit} = \left[\exp\left(\lambda_{i,t|t-1}\right)\right]^2 \tag{38}$$

following the formula (32); $\lambda_{i,t|t-1}$ is determined as a first order model defined by (34) or (35) respectively without or with a leverage effect. Moreover, the $\boldsymbol{\varepsilon}_t$ have a conditional multivariate t_ν–distribution with positive degrees of freedom ν:

$$\boldsymbol{\varepsilon}_t|\Omega_{t-1} \sim t_\nu\left(0, \mathbf{H}_t\right). \tag{39}$$

We use the same notations as in the DCC-GARCH model: $\mathbf{H}_t$ is given by (6) (i.e. $\mathbf{H}_t = \mathbf{S}_t\mathbf{P}_t\mathbf{S}_t$), the $N \times N$ conditional correlation matrix $\mathbf{P}_t$ is written as in (29), $\mathbf{S}_t$ has the following form:

$$\mathbf{S}_t = diag\left(\hat{\sigma}_{11,t}, \ldots, \hat{\sigma}_{NN,t}\right). \tag{40}$$

To precise the different diagonal elements of $\mathbf{S}_t$, we refer to Harvey and Chakravarty (2008): they defined $\sigma^*_{ii,t} = \sqrt{h_{iit}}$ as the time-varying scale parameter, which is not necessarily equal to the standard deviation. We resample the variance of residuals $\varepsilon^*_{i,t} \sim IID(0, \sigma^2_{\varepsilon^*})$ which does not necessarily have unit variance by the sample variance of residuals $\hat{\sigma}^2_{\varepsilon^*} = \sum \left(\hat{\varepsilon}^*_{i,t} - \bar{\hat{\varepsilon}}^*_i\right)^2 (T-1)^{-1}$, and we obtain thus the conditions which ensure the existence of a unit variance with $\hat{\sigma}_{ii,t} = \hat{\sigma}_{\varepsilon^*}\sigma^*_{ii,t}$. At last, $\mathbf{Q}_t = (q_{ij,t})$ is a symmetric positive definite $N \times N$ matrix given by (30) with $\bar{\mathbf{Q}}$ a $N \times N$ constant matrix.

The generation parameters are given in Table 7. Table 9 shows the empirical powers of the different tests, i.e. the rejection probabilities of the null hypothesis of constant correlations. For the $DCC - t - GARCH$ model, all the tests reject the null hypothesis of constancy of conditional correlation, but the $STCC$ and TSE tests present less good results and the $TAYLOR$ test the best one for each sample size. For the $DCC - Beta - t - EGARCH$ model with and without leverage effects, the performances of the tests present the same characteristics: the two new tests show the best empirical powers, even for $T = 1000$ or 1500, and the TSE test again seems the least powerful.

[**Insert Table 9 here**]

Everything considered, it is worthwhile noting here that the rejection probabilities of the null hypothesis increase again with the number of observations T. To conclude about the simulation results in the case of time-varying conditional correlation and non-normality assumptions, the tests showing the best performances are the $TAYLOR$ and $NEURAL$ tests. On the other hand, the tests presenting very poor performances are the $STCC$ and especially the

TSE tests.

5 Correlations between three asset returns

We illustrate the applicability of the two constant correlation tests that we propose, that is the ANN-based test and the Taylor-expansion based test, by comparing them to Tse's (2000) test and Silvennoinen and Teräsvirta's (2015) test in an empirical example. We consider the asset returns of three of "blue-chip" US daily stocks used to compute the Dow Jones Industrial Average index: JPMorgan Chase & Co. (JPM), The Coca-Cola Company (KO) and Exxon Mobil Corporation (XOM) from January 03, 2000 to October 12, 2012 ($T = 3217$ daily observations). The series correspond to closing prices adjusted for dividends and splits and are obtained from Yahoo Finance.

We compute the returns of these components of the Dow Jones index by the usual formula expressed for one day lag: $100 \times log(P_t/P_{t-1})$, where P_t represents the daily closing price at time t, $t = 1, \ldots, T$. To avoid a singularity issue of the Hessian matrix for the multivariate optimization and in order to have a comparable metric between the four tests, we removed the few peaks, caused by the market structure, by cutting them off at -10. Table 10 presents the summary statistics of the asset returns; medians, means, standard deviations, skewness and kurtosis take into account the cut off. The returns exhibit a positive excess of kurtosis.

[**Insert Table 10 here**]

We compute all the tests on the asset returns as bivariate and trivariate combinations in order to determine whether these series are conditionally correlated over time or not. Like in the simulation experiments, the number of hidden units is $p = 20$ and the number of lags of residuals is $q = 3$ for the test statistic $NEURAL$; the order of the Taylor expansion is $k = 3$ and the number of lags of residuals is $q = 2$ for the test statistic $TAYLOR$. Following Silvennoinen and Teräsvirta (2015), the $STCC$ test uses two different transition variables: the lagged absolute daily returns over seven days ($STCC_1$) and the contemporaneous squared daily returns ($STCC_2$), these returns being those of one or the other variable as appropriate. We choose in each case the transition variable maximizing the value of the statistic (or minimizing the corresponding $p - value$).

Table 11 shows the results for the different tests.

[**Insert Table 11 here**]

In the case of the relationship between JPMorgan Chase (JPM) and Coca-Cola returns (KO), the ANN-based test and the Taylor-expansion based test reject the null hypothesis of the constancy of conditional correlations at the 1% level (the $p - value$ are respectively 0.001 and 0.005) whereas TSE rejects it only at the 10% level (the $p - value$ is equal to 0.062). It is important to remark that $STCC_1$ and $STCC_2$ do not reject the null hypothesis considering that conditional correlations are not time-varying.

In the case of the JPMorgan Chase (JPM) and Exxon Mobil (XOM) returns, the $NEURAL$ test rejects the constant correlation hypothesis at the nominal size of 10% ($p - value = 0.067$) like the $TAYLOR$ test at the 1% level ($p - value = 8 \times 10^{-8}$), the TSE test at the 5% level ($p - value = 0.037$)

and $STCC_1$ at the 10% level ($p - value = 0.087$). $STCC_2$ does not reject the constancy of conditional correlations.

When we consider Exxon Mobil (XOM) and The Coca-Cola Company (KO) returns, the test based on artificial neural network rejects the hypothesis of constancy of conditional correlation at the nominal size of 5% ($p - value = 0.015$). The test based on a Taylor expansion rejects the null at the 1% level ($p - value = 3 \times 10^{-6}$) whereas the TSE and $STCC_1$ tests do not reject the constancy. $STCC_2$ rejects the hypothesis of constancy of conditional correlation at the nominal size of 5% ($p - value = 0.011$).

For the trivariate case (JPM-KO-XOM), all the tests reject the null hypothesis at the 1% level of significance, except for $STCC_2$ which rejects the constancy at the 5% level.

Therefore, if we had considered only the $STCC$ test in the first case, $STCC_2$ in the second case and TSE in the third case, we would have accepted the constancy of conditional correlations while the other tests conclude to time-varying conditional correlations. These short illustrations can thus show that the tests developed in this article have nice properties. This implies that whether we want to determine that the correlation is constant or not, both our tests can be considered as general misspecification tests.

6 Concluding remarks

In this paper, we propose two tests for the constancy of conditional correlations in the MGARCH models: the first one is based on artificial neural networks and the second on a Taylor expansion of each unknown conditional correlation around given point in a sample space. The main practical findings in this paper are that these two new tests perform well in our small-sample simulations and that they keep very good performances in case of non-normality: the simulations show their robustness to non-normality when the DGP models are multivariate versions of $GARCH-t$ and $Beta-t-EGARCH$ models. They show better performances with respect to the $STCC$ and TSE tests in most of cases probably because their original form is very general. Indeed, the first test is based on artificial neural networks that are universal approximators, and the second one comes from the linearization of an unknown relationship determining each conditional correlation.

In the empirical applications, both tests should be applied jointly because they have different finite-sample properties, as shown by the small sample simulations and the financial application. When we consider the powers, the $NEURAL$ test seems to have better properties when the variability of the conditional correlation coefficients is low; it is the opposite for the $TAYLOR$ test. Therefore, it is important to apply both of them jointly. The idea is to use a more complex model than a constant conditional correlations model if at least one of the two tests rejects the null hypothesis of constant conditional correlations.

We can thus conclude that they approximate quite well unknown specifications of the conditional correlations and therefore are useful in investigating potential time-varying conditional correlations between economic or financial time series.

APPENDIX: DERIVATION OF THE STATISTICS OF THE ARTIFICIAL NEURAL NETWORK TEST AND THE TAYLOR EXPANSION-BASED TEST

In order to compute the $NEURAL$ and the $TAYLOR$ statistics, we use broadly Abadir and Magnus (2005), Anderson (2003) and Lütkepohl (1996) for the derivations based on the rules of matrix algebra. Let us call r^* the number of parameters that are equal to zero under the null hypothesis, i.e. $r^* = p$ for the $NEURAL$ statistic and $r^* = N^*$ for the $TAYLOR$ statistic.

Each test statistic can be written (we simplify by T):

$$STATISTIC = \left(\sum_{t=1}^{T} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)' \Im(\boldsymbol{\theta})^{-1} \left(\sum_{t=1}^{T} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right). \tag{A.1}$$

The vector $\boldsymbol{\theta}$ of all the parameters of the model is given by: $\boldsymbol{\theta} = (\boldsymbol{\omega}', \boldsymbol{\rho}', \boldsymbol{\delta}')'$; $\boldsymbol{\omega} = (\omega_1', ..., \omega_N')$ is composed by the 3×1 parameter vector for the conditional variances $\boldsymbol{\omega}_i = (\zeta_i, \alpha_i, \beta_i)'$ for $i = 1, ..., N$, $\boldsymbol{\rho}$ is the vector of the conditional correlations ρ_{ij} with $1 \leq i < j \leq N$ and $\boldsymbol{\delta}$ is the vector of parameters that are equal to zero under the null hypothesis. Therefore, $\sum_{t=1}^{T} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ is a $(3N + (1 + r^*) \frac{N(N-1)}{2}) \times 1$ vector as follows

$$\sum_{t=1}^{T} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} V_\omega \\ V_\rho \\ V_\delta \end{pmatrix} \tag{A.2}$$

and $\Im(\boldsymbol{\theta})$ is a $(3N + (1 + r^*) \frac{N(N-1)}{2}) \times (3N + (1 + r^*) \frac{N(N-1)}{2})$ matrix given by

$$\Im(\boldsymbol{\theta}) = \begin{pmatrix} M_{\omega\omega} & M'_{\rho\omega} & M'_{\delta\omega} \\ M_{\rho\omega} & M_{\rho\rho} & M'_{\delta\rho} \\ M_{\delta\omega} & M_{\delta\rho} & M_{\delta\delta} \end{pmatrix}. \tag{A.3}$$

So we can write:

$$STATISTIC = \begin{pmatrix} V_\omega \\ V_\rho \\ V_\delta \end{pmatrix}' \begin{pmatrix} M_{\omega\omega} & M'_{\rho\omega} & M'_{\delta\omega} \\ M_{\rho\omega} & M_{\rho\rho} & M'_{\delta\rho} \\ M_{\delta\omega} & M_{\delta\rho} & M_{\delta\delta} \end{pmatrix}^{-1} \begin{pmatrix} V_\omega \\ V_\rho \\ V_\delta \end{pmatrix} \tag{A.4}$$

1. **Different sub-vectors of** $\sum_{t=1}^{T} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$

$$V_\omega = \begin{pmatrix} \sum_{t=1}^{T} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \omega'_1} & \cdots & \sum_{t=1}^{T} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \omega'_N} \end{pmatrix}', \; V_\rho = \left(\sum_{t=1}^{T} \frac{\partial l_t(\boldsymbol{\theta})}{\partial vecl(\mathbf{P})} \right), \tag{A.5}$$

and

$$V_\delta = vecl \left[\left(\sum_{t=1}^{T} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \delta_{ij}} \right)_{i,j} \right], \tag{A.6}$$

where V_ω, V_ρ and V_δ are respectively $3N \times 1$, $\frac{N(N-1)}{2} \times 1$ and $r^* \frac{N(N-1)}{2} \times 1$ vectors. *vecl* is an operator stacking the columns of the strict lower diagonal, i.e. excluding the diagonal elements of the matrix. Moreover, we have, for $i = 1, ..., N$ and $1 \leq i < j \leq N$:

$$\begin{cases} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \omega_i} = \frac{1}{2} \frac{1}{h_{iit}} \frac{\partial h_{iit}}{\partial \omega_i} \left[-1 + z_{it} \mathbf{1}'_i \mathbf{P}_t^{-1} \mathbf{z}_t \right] \\ \frac{\partial l_t(\boldsymbol{\theta})}{\partial vecl(\mathbf{P})} = -\frac{1}{2} \mathbf{U}' \left[vec\left(\mathbf{P}_t^{-1}\right) - \left(\mathbf{P}_t^{-1} \otimes \mathbf{P}_t^{-1}\right) (\mathbf{z}_t \otimes \mathbf{z}_t) \right] \\ \frac{\partial l_t(\boldsymbol{\theta})}{\partial \delta_{ij}} = \frac{\partial l_t(\boldsymbol{\theta})}{\partial \delta_{ji}} = -\frac{1}{2} W_{ijt} \left[vec\left(\mathbf{P}_t^{-1}\right) - \left(\mathbf{P}_t^{-1} \otimes \mathbf{P}_t^{-1}\right) (\mathbf{z}_t \otimes \mathbf{z}_t) \right] \end{cases} \tag{A.7}$$

where $\frac{\partial l_t(\boldsymbol{\theta})}{\partial \omega_i}$, $\frac{\partial l_t(\boldsymbol{\theta})}{\partial vecl(\mathbf{P})}$ and $\frac{\partial l_t(\boldsymbol{\theta})}{\partial \delta_{ij}}$ are respectively 3×1, $\frac{N(N-1)}{2}\times 1$ and $r^*\times 1$ vectors. Moreover, $\mathbf{z}_t = \mathbf{S}_t^{-1}\boldsymbol{\varepsilon}_t$ and the matrix $\mathbf{U}$ is an $N^2\times\frac{N(N-1)}{2}$ matrix of zeros and ones whose columns are defined as

$$\left(vec\left(1_i 1_j' + 1_j 1_i'\right)\right)_{\substack{i=1,\dots,N-1\\ j=i+1,\dots,N}} \tag{A.8}$$

and the columns appear in the same order from left to right as the indices in $vecl(\mathbf{P})$. The matrix W_{ijt} is an $r^*\times N^2$ matrix given by

$$W_{ijt} = diag\left(\mathbf{g}_{ijt}\right)\left\{\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}\otimes\left(vec\left(1_i 1_j' + 1_j 1_i'\right)\right)'\right\}; \tag{A.9}$$

$\mathbf{g}_{ijt}$ is a $r^*\times 1$ vector defined in the $NEURAL$ test as

$$\mathbf{g}_{ijt} = \begin{pmatrix} g_{ijt,1}\\ \vdots\\ g_{ijt,r^*}\end{pmatrix} = \begin{pmatrix}(1+\exp\{-w_{ijt}'\gamma_{ij1}\})^{-1}\\ \vdots\\ (1+\exp\{-w_{ijt}'\gamma_{ijr^*}\})^{-1}\end{pmatrix} \tag{A.10}$$

and, in the $TAYLOR$ test, it is replaced by $\mathbf{d}_{ijt}$, composed by all terms involving functions of elements of lagged values of $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$; we have:

$$r^* = 2\left[\sum_{r=1}^{k}\begin{pmatrix}2q+r-1\\ r\end{pmatrix} - \sum_{r=1}^{k}\begin{pmatrix}q+r-1\\ r\end{pmatrix}\right]. \tag{A.11}$$

2. **Different elements of the matrix** $\Im(\theta)$

- $M_{\omega\omega}$ is the following $3N \times 3N$ matrix:

$$M_{\omega\omega} = \sum_{t=1}^{T} \begin{pmatrix} M_{\omega\omega,11} & M'_{\omega\omega,12} & \cdots & M'_{\omega\omega,1N} \\ M_{\omega\omega,12} & M_{\omega\omega,22} & \cdots & M'_{\omega\omega,2N} \\ \vdots & \vdots & & \vdots \\ M_{\omega\omega,1N} & M_{\omega\omega,2N} & \cdots & M_{\omega\omega,NN} \end{pmatrix} \tag{A.12}$$

where, for $i = 1, ..., N$ and $1 \leq i < j \leq N$, $M_{\omega\omega,ii}$ and $M_{\omega\omega,ij}$ are 3×3 matrices defined as

$$\begin{cases} M_{\omega\omega,ii} = E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta})}{\partial \omega_i}\frac{\partial l_t(\boldsymbol{\theta})}{\partial \omega_i'}\right] = \frac{1}{4}\frac{1}{h_{iit}^2}\frac{\partial h_{iit}}{\partial \omega_i}\left[\mathbf{1}_i'\mathbf{P}_t^{-1}\mathbf{1}_i + \mathbf{1}\right]\frac{\partial h_{iit}}{\partial \omega_i}' \\ M_{\omega\omega,ij} = E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta})}{\partial \omega_i}\frac{\partial l_t(\boldsymbol{\theta})}{\partial \omega_j'}\right] = \frac{1}{4}\frac{1}{h_{iit}h_{jjt}}\frac{\partial h_{iit}}{\partial \omega_i}\left[\rho_{ijt}\mathbf{1}_i'\mathbf{P}_t^{-1}\mathbf{1}_j + \mathbf{1}\right]\frac{\partial h_{jjt}}{\partial \omega_j}' \end{cases} \tag{A.13}$$

- $M_{\rho\rho}$ is the following $\frac{N(N-1)}{2} \times \frac{N(N-1)}{2}$ matrix:

$$M_{\rho\rho} = \sum_{t=1}^{T}\frac{1}{4}\mathbf{U}'\left[\left(\mathbf{P}_t^{-1}\otimes\mathbf{P}_t^{-1}\right) + \left(\mathbf{P}_t^{-1}\otimes\mathbf{I}\right)\mathbf{K}(\mathbf{P}_t^{-1}\otimes\mathbf{I})\right]\mathbf{U} \tag{A.14}$$

- $M_{\rho\omega}$ is a $\frac{N(N-1)}{2} \times 3N$ matrix defined as

$$M_{\rho\omega} = \sum_{t=1}^{T}\begin{pmatrix} M'_{\rho\omega,1} & M'_{\rho\omega,2} & ... & M'_{\rho\omega,N} \end{pmatrix} \tag{A.15}$$

where each $3 \times \frac{N(N-1)}{2}$ matrix $M_{\rho\omega,i}$, for $i = 1, ..., N$, is:

$$\begin{aligned} M_{\rho\omega,i} &= E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta})}{\partial \omega_i}\frac{\partial l_t(\boldsymbol{\theta})}{\partial vecl(\mathbf{P})'}\right] \\ &= \frac{1}{4}\frac{1}{h_{iit}}\frac{\partial h_{iit}}{\partial \omega_i}\left[\mathbf{1}_i'\left(\mathbf{1}_i'\otimes\mathbf{P}_t^{-1}\right) + \mathbf{1}_i'\left(\mathbf{1}_i'\otimes\mathbf{I}\right)\mathbf{K}(\mathbf{P}_t^{-1}\otimes\mathbf{I})\right]\mathbf{U} \end{aligned} \tag{A.16}$$

- $M_{\delta\rho}$ is a $r^*\frac{N(N-1)}{2} \times \frac{N(N-1)}{2}$ matrix given by

$$M_{\delta\rho} = \sum_{t=1}^{T} vecl\left[(M_{\delta\rho,ij}\ )_{i,j}\right] \tag{A.17}$$

where each $r^* \times \frac{N(N-1)}{2}$ matrix $M_{\delta\rho,ij}$ is

$$\begin{aligned} M_{\delta\rho,ij} = M_{\delta\rho,ji} = E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta})}{\partial \delta_{ij}} \frac{\partial l_t(\boldsymbol{\theta})}{\partial vecl(\mathbf{P})'}\right] \\ = \tfrac{1}{4} W_{ijt}\left[\left(\mathbf{P}_t^{-1} \otimes \mathbf{P}_t^{-1}\right) + \left(\mathbf{P}_t^{-1} \otimes \mathbf{I}\right) \mathbf{K} \left(\mathbf{P}_t^{-1} \otimes \mathbf{I}\right)\right] \mathbf{U} \end{aligned} \tag{A.18}$$

- $M_{\delta\delta}$ is the following $r^*\frac{N(N-1)}{2} \times r^*\frac{N(N-1)}{2}$ matrix:

$$M_{\delta\delta} = \sum_{t=1}^{T} \left[M_{\delta\delta}(i,j,n,m)\right]_{i,j,n,m} \tag{A.19}$$

with $M_{\delta\delta}(i,j,n,m)$ a $r^* \times r^*$ matrix given by:

$$\begin{aligned} M_{\delta\delta}(i,j,n,m) = E_{t-1}\left[\frac{\partial l_t(\boldsymbol{\theta})}{\partial \delta_{ij}} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \delta'_{nm}}\right] \\ = \tfrac{1}{4} W_{ijt}\left[\left(\mathbf{P}_t^{-1} \otimes \mathbf{P}_t^{-1}\right) + \left(\mathbf{P}_t^{-1} \otimes \mathbf{I}\right) \mathbf{K}(\mathbf{P}_t^{-1} \otimes \mathbf{I})\right] W'_{nmt}; \end{aligned} \tag{A.20}$$

for each couple (i,j) taken in the same order as for V_δ, take the couple (n,m) in the same order.

- $M_{\delta\omega}$ is a $r^*\frac{N(N-1)}{2} \times 3N$ matrix given by:

$$M_{\delta\omega} = \sum_{t=1}^{T} \left[M_{\delta\omega}\ (i,j,n)\right]_{i,j,n} \tag{A.21}$$

with $M_{\delta\omega}(i,j,n)$ a $r^* \times 3$ matrix defined by:

$$
\begin{aligned}
M_{\delta\omega}(i,j,n) &= E_{t-1}\left[\tfrac{\partial l_t(\boldsymbol{\theta})}{\partial \delta_{ij}}\tfrac{\partial l_t(\boldsymbol{\theta})}{\partial \omega_n'}\right] \\
&= -\tfrac{1}{4h_{nnt}} W_{ijt}\left[vec(\mathbf{P}_t^{-1}) - \left\{(\mathbf{I}\otimes \mathbf{1}_n) + \left(\mathbf{P}_t^{-1}\otimes \mathbf{I}\right)\mathbf{K}(\mathbf{I}\otimes \mathbf{P}_t\mathbf{1}_n)\right.\right. \\
&\quad \left.\left. + vec(\mathbf{P}_t^{-1})vec(\mathbf{1}_n'\mathbf{P}_t)'\right\}\mathbf{P}_t^{-1}\mathbf{1}_n\right]\tfrac{\partial h_{nnt}}{\partial \omega_n'};
\end{aligned} \tag{A.22}
$$

for each couple (i,j) taken in the same order as for V_δ, take all the $n = 1, 2, ..., N$.

For the computation of all these vectors and matrices, we have:

$$
\frac{\partial h_{iit}}{\partial \omega_i} = \nu_{i,t-1} + \beta_i \frac{\partial h_{ii,t-1}}{\partial \omega_i}. \tag{A.23}
$$

with $\nu_{it} = \begin{pmatrix} 1, & \varepsilon_{it}^2, & h_{iit} \end{pmatrix}'$ for $i = 1, ..., N$ (we compute recursively $\frac{\partial h_{iit}}{\partial \omega_i}$).

Following Silvennoinen and Teräsvirta (2015), we have for a model with general correlation matrix $\mathbf{P}_t$

$$
\begin{cases}
E_{t-1}\left[\mathbf{z}_t\mathbf{z}_t' \otimes \mathbf{z}_t\mathbf{z}_t'\right] = (\mathbf{P}_t \otimes \mathbf{P}_t) + (\mathbf{I}\otimes \mathbf{P}_t)\mathbf{K}(\mathbf{I}\otimes \mathbf{P}_t) + vec\mathbf{P}_t(vec\mathbf{P}_t)' \\
E_{t-1}(z_{it}z_{jt}' \otimes \mathbf{z}_t\mathbf{z}_t') = (\mathbf{1}_i' \otimes \mathbf{I})\, E_{t-1}(\mathbf{z}_t\mathbf{z}_t' \otimes \mathbf{z}_t\mathbf{z}_t')\,(\mathbf{1}_j \otimes \mathbf{I}) \\
E_{t-1}\left[z_{it}\mathbf{z}_t' \otimes \mathbf{z}_t\mathbf{z}_t'\right] = (\mathbf{1}_i' \otimes \mathbf{I})\, E_{t-1}\left[\mathbf{z}_t\mathbf{z}_t' \otimes \mathbf{z}_t\mathbf{z}_t'\right]
\end{cases} \tag{A.24}
$$

with $\mathbf{K}$ a $N^2 \times N^2$ matrix defined as

$$\mathbf{K} = \begin{bmatrix} 1_1 1_1' & \cdots & 1_i 1_1' & \cdots & 1_N 1_1' \\ \vdots & & \vdots & & \vdots \\ 1_1 1_j' & \cdots & 1_i 1_j' & \cdots & 1_N 1_j' \\ \vdots & & \vdots & & \vdots \\ 1_1 1_N' & \cdots & 1_i 1_N' & \cdots & 1_N 1_N' \end{bmatrix} \tag{A.25}$$

Following (A.4), and using the rules of the partitioned matrices and the null hypothesis, the test statistic can be written

$$STATISTIC = V_\delta' \left[M_{\delta\delta} - \begin{pmatrix} M_{\delta\omega} & M_{\delta\rho} \end{pmatrix} \begin{pmatrix} M_{\omega\omega} & M_{\rho\omega}' \\ M_{\rho\omega} & M_{\rho\rho} \end{pmatrix}^{-1} \right. \tag{A.26}$$

$$\left. \times \begin{pmatrix} M_{\delta\omega}' \\ M_{\delta\rho}' \end{pmatrix} \right]^{-1} V_\delta .$$

References

[1] Abadir, K.M., Magnus, J.R. (2005). *Matrix Algebra*. Cambridge University Press.

[2] Aielli, G.P. (2013). Dynamic Conditional Correlation: on Properties and Estimation. *Journal of Business and Economic Statistics* 31: 282-299.

[3] Anderson, T. W. (2003), *An Introduction to Multivariate Statistical Analysis*. New York: Wiley, 3 edn.

[4] Baillie, R.T., Bollerslev, T. (1990). A multivariate generalized ARCH approach to modeling risk premia in forward foreign exchange rate markets. *Journal of International Money and Finance* 9:309-324.

[5] Bauwens, L., Laurent, S., Rombouts, J.V.K. (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21:79–109.

[6] Bera, AK, Kim, S. (2002). Testing constancy of correlation and other specifications of the BGARCH model with an application to international equity returns. *Journal of Empirical Finance* 9:171–195.

[7] Berben, R.P., Jansen, W.J. (2005). Comovement in international equity markets: A sectoral view. *Journal of International Money and Finance* 24:832-857.

[8] Bollerslev, T. (1987). A conditionally heteroskedastic time series model for security prices and rates of return data. *Review of Economics and Statistics* 59:542-547.

[9] Bollerslev, T. (1990). Modelling the coherence in short run nominal exchange rates: a multivariate generalized ARCH model. *Review of Economics and Statistics* 72:498-505.

[10] Carroll, S.M., Dickinson, B.W. (1989). Construction of neural nets using the Radon transform. *Proceedings of the international joint conference on neural networks*, Washington, D.C., IEEE Press I, 607-611.

[11] Castle, J. L., Hendry, D. F. (2010). A low-dimension portmanteau test for non-linearity. *Journal of Econometrics* 150:231-245.

[12] Caulet, R., Péguin-Feissolle, A. (2000). Un test d'hétéroscédasticité conditionnelle inspiré de la modélisation en termes de réseaux neuronaux artificiels. *Annales d'Économie et de Statistique* 59:177-197.

[13] Cybenko, G. (1989). Approximation by superpositions of a sigmoid function. *Mathematics of Control, Signals and Systems* 2:303-314.

[14] Ding, Z., Engle, R.F. (2001). Large scale conditional covariance matrix modeling, estimation and testing. *Academia Economic Papers* 29:157-184.

[15] Engle, R.F. (2002). Dynamic conditional correlation—a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics* 20:339–350.

[16] Engle, R.F., Kroner F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory* 11:122-150.

[17] Engle, RF, Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. NBER Working Paper 8554.

[18] Harvey, A., Chakravarty, T. (2008). Beta-t-(E)GARCH. Cambridge Working Papers in Economics, CWPE 0840.

[19] Harvey, A., Sucarrat, G. (2014). EGARCH models with fat tails, skewness and leverage. *Computational Statistics and Data Analysis* 76: 320-338.

[20] Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. *Neural Networks* 4:251-257.

[21] Hornik, K., Stinchcombe, M., White, H. (1989). Multi-layer feedforward networks are universal approximators. *Neural Networks* 2:359-366.

[22] Hornik, K., Stinchcombe, M., White, H. (1990). Universal approximation of an unknown mapping and its derivatives using multi-layer feedforward networks. *Neural Networks* 3:551-560.

[23] Jeantheau, T. (1998). Strong consistency of estimators for multivariate ARCH models. *Econometric Theory* 14:70-86.

[24] Kamstra, M. (1993). A neural network test for heteroscedasticity. Working Paper, Simon Fraser University, Burnaby, Canada.

[25] Lebreton, M., Péguin-Feissolle, A. (2007). Robust tests for heteroscedasticity in a general framework. *Annales d'Économie et de Statistique* 85:159-187.

[26] Lee, T.H., White, H., Granger, C.W.J. (1993). Testing for neglecting nonlinearity in time series models. *Journal of Econometrics* 56:269-290.

[27] Longin, F., Solnik, B. (1995). Is the correlation in international equity returns constant: 1960-1990?. *Journal of International Money and Finance* 14:3-26.

[28] Lütkepohl, H. (1996). *Handbook of Matrices.* Chichester: John Wiley & Sons.

[29] Orme, C. (1990). The small sample performance of the information matrix test. *Journal of Econometrics* 46:309–331.

[30] Péguin-Feissolle, A. (1999). A comparison of the power of some tests for conditional heteroscedasticity. *Economics Letters* 63:5-17.

[31] Péguin-Feissolle, A., Strikholm, B., Teräsvirta, T. (2013). Testing the Granger noncausality hypothesis in stationary nonlinear models of unknown functional form. *Communications in Statistics - Simulation and Computation* 42:1063-1087.

[32] Péguin-Feissolle, A., Teräsvirta, T. (1999). A general framework for testing the Granger noncausality hypothesis. SSE/EFI Working Paper Series in Economics and Finance n° 343, Stockholm School of Economics.

[33] Silvennoinen, A., Teräsvirta, T. (2009a). Multivariate GARCH Models. In *Handbook of Financial Time Series*, eds. T. G. Andersen, R. A. Davis, J.-P. Kreiss, T. Mikosch, 201–229, New York: Springer.

[34] Silvennoinen, A., Teräsvirta, T. (2009b). Modeling Multivariate Autoregressive Conditional Heteroskedasticity with the Double Smooth Transition Conditional Correlation GARCH Model. *Journal of Financial Econometrics* 7:373–411.

[35] Silvennoinen, A., Teräsvirta, T.(2015). Modelling conditional correlations of asset returns: A smooth transition approach. *Econometric Reviews* 34:174-197.

[36] Stinchcombe, M., White, H. (1989). Universal approximation using feedforward networks with non-sigmoid hidden layer activation functions. In *Proceedings of the international joint conference on neural networks*, Washington, D.C., IEEE Press, I, 613-618.

[37] Teräsvirta, T., Lin, C.F., Granger, C.W.J. (1993). Power of the neural network linearity test. *Journal of Time Series Analysis* 14:209-220.

[38] Tse, Y. K. (2000). A test for constant correlations in a multivariate GARCH model. *Journal of Econometrics* 98:107–127.

[39] White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica* 50:1–25.

Table 1: Determination of the optimal order of the Taylor expansion k and the optimal number of lags of residuals q used in the ANN-based and Taylor expansion-based tests (ECCC-GARCH)

T	**1000**			**1500**			**2500**		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
TAYLOR k = 1									
q = 1	0.009	0.050	0.113	0.012	0.043	0.089	0.008	0.039	0.094
q = 2	0.016	0.063	0.123	0.013	0.055	0.097	0.006	0.059	0.124
q = 3	0.023	0.086	0.144	0.012	0.042	0.087	0.010	0.046	0.102
q = 4	0.017	0.070	0.131	0.010	0.059	0.110	0.006	0.044	0.088
TAYLOR k = 2									
q = 1	0.083	0.159	0.225	0.081	0.166	0.231	0.007	0.052	0.102
q = 2	0.005	0.028	0.061	0.008	0.031	0.073	0.008	0.051	0.105
q = 3	0.022	0.058	0.122	0.006	0.035	0.079	0.012	0.060	0.114
q = 4	0.020	0.068	0.114	0.008	0.049	0.103	0.011	0.063	0.131
TAYLOR k = 3									
q = 1	0.032	0.089	0.148	0.037	0.098	0.150	0.005	0.062	0.109
q = 2	0.012	0.042	0.083	0.014	0.053	0.097	0.007	0.049	0.106
q = 3	0.032	0.096	0.144	0.011	0.065	0.131	0.018	0.056	0.119
q = 4	0.024	0.086	0.151	0.022	0.093	0.155	0.021	0.067	0.126
TAYLOR k = 4									
q = 1	0.049	0.095	0.133	0.036	0.073	0.111	0.009	0.056	0.104
q = 2	0.014	0.045	0.090	0.009	0.050	0.096	0.011	0.058	0.108
q = 3	0.026	0.083	0.146	0.008	0.056	0.117	0.016	0.061	0.122
q = 4	0.024	0.080	0.160	0.024	0.094	0.146	0.018	0.069	0.122
NEURAL									
q = 1	0.004	0.018	0.035	0.003	0.018	0.039	0.006	0.058	0.109
q = 2	0.009	0.027	0.055	0.007	0.026	0.057	0.013	0.058	0.107
q = 3	0.007	0.052	0.086	0.008	0.048	0.083	0.009	0.052	0.103
q = 4	0.009	0.036	0.080	0.007	0.035	0.072	0.009	0.052	0.108

Note: The number of replications is $S = 1000$. The generation parameters of the ECCC model are:

ζ_1	ζ_2	α_{11}	α_{12}	α_{21}	α_{22}	β_{11}	β_{12}	β_{21}	β_{22}	ρ_{12}
0.02	0.02	0.01	0.02	0.02	0.01	0.85	0.02	0.02	0.85	0.60

Table 2: Determination of the optimal order of the Taylor expansion k and the optimal number of lags of residuals q used in the ANN-based and Taylor expansion-based tests (GARCH)

T		**1000**			**1500**			**2500**	
	1%	**5%**	**10%**	**1%**	**5%**	**10%**	**1%**	**5%**	**10%**
TAYLOR k = 1									
q = 1	0.015	0.062	0.108	0.007	0.048	0.103	0.008	0.042	0.091
q = 2	0.018	0.067	0.110	0.012	0.045	0.098	0.012	0.057	0.118
q = 3	0.018	0.066	0.117	0.014	0.051	0.100	0.014	0.056	0.098
q = 4	0.016	0.062	0.130	0.011	0.049	0.105	0.008	0.049	0.106
TAYLOR k = 2									
q = 1	0.010	0.049	0.085	0.006	0.038	0.088	0.011	0.055	0.095
q = 2	0.010	0.052	0.099	0.011	0.051	0.097	0.011	0.046	0.101
q = 3	0.013	0.062	0.111	0.005	0.045	0.098	0.017	0.058	0.111
q = 4	0.016	0.062	0.118	0.006	0.050	0.106	0.014	0.067	0.124
TAYLOR k = 3									
q = 1	0.013	0.058	0.112	0.007	0.042	0.091	0.007	0.045	0.103
q = 2	0.011	0.041	0.091	0.011	0.052	0.101	0.011	0.052	0.109
q = 3	0.015	0.051	0.105	0.006	0.051	0.101	0.012	0.053	0.114
q = 4	0.017	0.068	0.119	0.014	0.044	0.092	0.010	0.046	0.093
TAYLOR k = 4									
q = 1	0.011	0.041	0.077	0.010	0.050	0.095	0.012	0.046	0.119
q = 2	0.007	0.043	0.079	0.012	0.060	0.104	0.009	0.044	0.084
q = 3	0.014	0.058	0.092	0.009	0.042	0.088	0.015	0.056	0.105
q = 4	0.022	0.075	0.130	0.012	0.052	0.103	0.013	0.062	0.118
NEURAL									
q = 1	0.010	0.056	0.112	0.012	0.048	0.092	0.012	0.048	0.104
q = 2	0.013	0.047	0.097	0.014	0.053	0.107	0.011	0.059	0.103
q = 3	0.011	0.047	0.089	0.006	0.045	0.088	0.010	0.051	0.106
q = 4	0.015	0.047	0.093	0.014	0.053	0.105	0.010	0.044	0.100

Note: The number of replications is $S = 1000$. The generation parameters of the bivariate GARCH model are:

ζ_1	ζ_2	α_{11}	α_{22}	β_{11}	β_{22}
0.02	0.01	0.04	0.03	0.95	0.96

Table 3: Small sample sizes of the different constant conditional correlation tests

T	1000			1500			2500		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
ECCC									
$NEURAL$	0.008	0.052	0.087	0.009	0.049	0.085	0.011	0.054	0.105
$TAYLOR$	0.012	0.042	0.083	0.014	0.053	0.097	0.009	0.053	0.104
TSE	0.027	0.084	0.134	0.023	0.080	0.139	0.086	0.144	0.201
$STCC$	0.011	0.054	0.107	0.012	0.042	0.094	0.011	0.049	0.104
CCC									
$NEURAL$	0.018	0.077	0.134	0.017	0.061	0.118	0.008	0.056	0.104
$TAYLOR$	0.021	0.083	0.141	0.017	0.070	0.133	0.010	0.046	0.105
TSE	0.018	0.060	0.117	0.015	0.062	0.107	0.017	0.052	0.105
$STCC$	0.015	0.059	0.112	0.013	0.059	0.106	0.008	0.047	0.098
GARCH									
$NEURAL$	0.014	0.053	0.107	0.015	0.057	0.113	0.011	0.055	0.106
$TAYLOR$	0.013	0.053	0.105	0.015	0.056	0.103	0.013	0.058	0.111
TSE	0.093	0.193	0.264	0.089	0.189	0.265	0.089	0.192	0.264
$STCC$	0.006	0.045	0.101	0.011	0.046	0.095	0.008	0.054	0.105
GJR									
$NEURAL$	0.013	0.061	0.119	0.008	0.052	0.099	0.010	0.061	0.109
$TAYLOR$	0.015	0.065	0.123	0.010	0.043	0.103	0.014	0.052	0.108
TSE	0.033	0.099	0.153	0.030	0.088	0.148	0.028	0.091	0.151
$STCC$	0.008	0.043	0.082	0.007	0.050	0.104	0.008	0.049	0.098

Note: The number of replications is $S = 2000$. For $NEURAL$ the number of hidden units is $p = 20$ and the number of lags of residuals is $q = 3$, and for $TAYLOR$ the order of the Taylor expansion is $k = 3$ and the number of lags of residuals is $q = 2$. The generation parameters are:

	ζ_1	ζ_2	α_{11}	α_{12}	α_{21}	α_{22}	β_{11}	β_{12}	β_{21}	β_{22}	γ_{11}	γ_{22}	ρ_{12}
ECCC	0.02	0.02	0.01	0.02	0.02	0.01	0.85	0.02	0.02	0.85	–	–	0.60
CCC	0.02	0.02	0.01	0.00	0.00	0.01	0.98	0.00	0.00	0.98	–	–	0.30
GARCH	0.40	0.20	0.15	0.00	0.00	0.20	0.80	0.00	0.00	0.70	–	–	0.00
GJR	0.02	0.01	0.05	0.00	0.00	0.04	0.91	0.00	0.00	0.93	0.02	0.01	0.00

Table 4: MGARCH parameters for the generation of the DGP for the small sample powers

	$\mathbf{BEKK_1}$	$\mathbf{BEKK_2}$	$\mathbf{DCC_1}$	$\mathbf{DCC_2}$	$\mathbf{EDCC_1}$	$\mathbf{EDCC_2}$	**ASBEKK**
ζ_{11}	0.20	0.20	0.01	0.01	0.01	0.01	0.7
ζ_{12}	0.01	0.04	–	–	–	–	0.3
ζ_{22}	0.20	0.20	0.02	0.02	0.02	0.02	0.9
α_{11}	0.10	0.30	0.001	0.001	0.01	0.02	–
α_{12}	0.10	0.10	0.00	0.00	0.01	0.05	–
α_{21}	0.10	0.10	0.00	0.00	0.01	0.05	–
α_{22}	0.10	0.30	0.002	0.002	0.02	0.02	–
β_{11}	0.70	0.30	0.98	0.98	0.80	0.75	–
β_{12}	0.10	0.20	0.00	0.00	0.15	0.10	–
β_{21}	0.10	0.20	0.00	0.00	0.15	0.10	–
β_{22}	0.70	0.30	0.90	0.90	0.75	0.55	–
α	–	–	0.05	0.15	0.15	0.01	0.02
β	–	–	0.94	0.80	0.84	0.94	0.90
γ	–	–	–	–	–	–	0.05
ρ_{12}	–	–	0.80	0.30	0.80	0.30	–
$\mathbf{corr}_{min_{1000}}$	0.33	0.18	0.26	−0.66	−0.65	0.21	0.49
$\mathbf{corr}_{max_{1000}}$	0.53	0.77	0.94	0.86	0.99	0.38	0.75
$\mathbf{corr}_{min_{1500}}$	0.33	0.17	0.19	−0.69	−0.75	0.21	0.49
$\mathbf{corr}_{max_{1500}}$	0.54	0.78	0.94	0.87	0.99	0.39	0.76
$\mathbf{corr}_{min_{2500}}$	0.33	0.15	0.11	−0.71	−0.81	0.20	0.49
$\mathbf{corr}_{max_{2500}}$	0.56	0.80	0.95	0.88	0.99	0.39	0.77

Table 5: Small sample powers of the different constant conditional correlation tests

T	1000			1500			2500		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathbf{BEKK_1}$									
$NEURAL$	0.998	0.998	0.998	1.000	1.000	1.000	1.000	1.000	1.000
$TAYLOR$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
TSE	0.832	0.986	0.995	0.989	0.998	0.999	1.000	1.000	1.000
$STCC$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\mathbf{BEKK_2}$									
$NEURAL$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$TAYLOR$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
TSE	0.983	0.987	0.990	0.995	1.000	1.000	1.000	1.000	1.000
$STCC$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ASBEKK									
$NEURAL$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$TAYLOR$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
TSE	0.406	0.578	0.666	0.5335	0.720	0.800	0.739	0.876	0.919
$STCC$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\mathbf{DCC_1}$									
$NEURAL$	0.392	0.536	0.612	0.410	0.546	0.619	0.433	0.585	0.659
$TAYLOR$	0.394	0.534	0.616	0.446	0.602	0.694	0.525	0.648	0.722
TSE	0.060	0.149	0.241	0.100	0.234	0.344	0.191	0.375	0.487
$STCC$	0.197	0.329	0.416	0.213	0.350	0.442	0.241	0.383	0.466
$\mathbf{DCC_2}$									
$NEURAL$	0.144	0.265	0.359	0.163	0.303	0.397	0.243	0.409	0.501
$TAYLOR$	0.970	0.988	0.992	0.999	0.999	0.999	1.000	1.000	1.000
TSE	0.829	0.943	0.966	0.967	0.991	0.997	0.997	0.999	1.000
$STCC$	0.237	0.369	0.454	0.232	0.388	0.463	0.249	0.397	0.456
$\mathbf{EDCC_1}$									
$NEURAL$	0.693	0.784	0.830	0.711	0.797	0.835	0.756	0.832	0.870
$TAYLOR$	0.923	0.956	0.969	0.969	0.981	0.986	0.994	0.996	0.996
TSE	0.799	0.865	0.895	0.874	0.909	0.920	0.911	0.917	0.929
$STCC$	0.454	0.566	0.631	0.509	0.604	0.654	0.615	0.708	0.747
$\mathbf{EDCC_2}$									
$NEURAL$	0.410	0.512	0.573	0.418	0.527	0.593	0.503	0.605	0.676
$TAYLOR$	0.023	0.088	0.163	0.031	0.104	0.178	0.127	0.188	0.247
TSE	0.022	0.081	0.138	0.025	0.094	0.158	0.131	0.199	0.262
$STCC$	0.064	0.156	0.238	0.069	0.165	0.243	0.154	0.250	0.335

Note: The number of replications is $S = 2000$. For $NEURAL$ the number of hidden units is $p = 20$ and the number of lags of residuals is $q = 3$, and for $TAYLOR$ the order of the Taylor expansion is $k = 3$ and the number of lags of residuals is $q = 2$.

Table 6: Small sample sizes and powers of the different constant conditional correlation tests when the number of endogenous variables is N = 3

T	**1000**			**1500**			**2500**		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathbf{GARCH_2}$									
NEURAL	0.008	0.039	0.086	0.013	0.047	0.091	0.010	0.046	0.082
TAYLOR	0.011	0.048	0.089	0.010	0.048	0.089	0.008	0.045	0.085
TSE	0.015	0.050	0.093	0.017	0.049	0.086	0.012	0.053	0.096
STCC	0.013	0.052	0.104	0.009	0.049	0.098	0.011	0.054	0.104
$\mathbf{BEKK_3}$									
NEURAL	0.993	0.996	0.996	0.996	0.996	0.996	0.999	0.999	0.999
TAYLOR	0.994	0.994	0.995	0.995	0.996	0.996	0.996	0.996	0.996
TSE	0.344	0.504	0.590	0.432	0.609	0.687	0.615	0.765	0.836
STCC	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note: The number of replications is $S = 2000$. For $NEURAL$ the number of hidden units is $p = 20$ and the number of lags of residuals is $q = 3$, and for $TAYLOR$ the order of the Taylor expansion is $k = 3$ and the number of lags of residuals is $q = 2$. The generation parameters are:

$\mathbf{GARCH_2}$

$$\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} = \begin{bmatrix} 0.04 \\ 0.02 \\ 0.01 \end{bmatrix} \qquad \begin{bmatrix} \alpha_{11} \\ \alpha_{22} \\ \alpha_{33} \end{bmatrix} = \begin{bmatrix} 0.05 \\ 0.04 \\ 0.03 \end{bmatrix} \qquad \begin{bmatrix} \beta_{11} \\ \beta_{22} \\ \beta_{33} \end{bmatrix} = \begin{bmatrix} 0.91 \\ 0.92 \\ 0.93 \end{bmatrix}$$

$\mathbf{BEKK_3}$

$$\mathbf{Z} = \begin{bmatrix} 0.80 & 0.20 & 0.20 \\ 0.00 & 0.80 & 0.20 \\ 0.00 & 0.00 & 0.80 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} 0.01 & 0.01 & 0.01 \\ 0.01 & 0.02 & 0.01 \\ 0.01 & 0.01 & 0.03 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0.80 & 0.02 & 0.02 \\ 0.02 & 0.80 & 0.02 \\ 0.02 & 0.02 & 0.80 \end{bmatrix}$$

Table 7: MGARCH parameters for the generation of the DGP for the small sample sizes and powers under nonnormality

	t-GARCH	**Beta-t-EGARCH**	**DCC-t-GARCH**	**DCC-Beta-t-EGARCH**
α	–	–	0.05	0.15
β	–	–	0.94	0.80
ρ_{12}	0.00	0.00	0.70	0.60
ν	10	10	10	10
ζ_1	0.40	–	0.30	–
ζ_2	0.20	–	0.20	–
α_{11}	0.15	–	0.10	–
α_{12}	0.00	–	0.00	–
α_{21}	0.00	–	0.00	–
α_{22}	0.20	–	0.20	–
β_{11}	0.80	–	0.70	–
β_{12}	0.00	–	0.00	–
β_{21}	0.00	–	0.00	–
β_{22}	0.70	–	0.60	–
μ_1	–	0.00	–	0.00
μ_2	–	0.00	–	0.00
δ_1	–	0.007	–	0.02
δ_2	–	0.05	–	0.04
ϕ_1	–	0.99	–	0.95
ϕ_2	–	0.90	–	0.90
κ_1	–	0.05	–	0.05
κ_2	–	0.07	–	0.07
κ_1^*	–	0.02* or 0.00	–	0.04* or 0.00
κ_2^*	–	0.04* or 0.00	–	0.06* or 0.00

*Note: * with leverage effect in the Beta-t-EGARCH and DCC-Beta-t-EGARCH*

Table 8: Small sample sizes of constant conditional correlation tests under nonnormality

T	**1000**			**1500**			**2500**		
	1%	**5%**	**10%**	**1%**	**5%**	**10%**	**1%**	**5%**	**10%**
t − GARCH									
NEURAL	0.014	0.058	0.116	0.016	0.059	0.102	0.016	0.056	0.102
TAYLOR	0.014	0.064	0.115	0.013	0.061	0.118	0.022	0.072	0.118
TSE	0.056	0.112	0.168	0.056	0.125	0.191	0.075	0.153	0.219
STCC	0.012	0.055	0.097	0.011	0.043	0.091	0.012	0.051	0.101
Beta − t − EGARCH									
NEURAL	0.024	0.057	0.107	0.017	0.064	0.117	0.014	0.053	0.100
TAYLOR	0.026	0.070	0.116	0.018	0.071	0.125	0.014	0.056	0.100
TSE	0.088	0.154	0.215	0.072	0.156	0.220	0.077	0.152	0.220
STCC	0.015	0.053	0.101	0.017	0.061	0.110	0.011	0.049	0.107
Beta − t − EGARCH*									
NEURAL	0.018	0.052	0.105	0.017	0.057	0.114	0.015	0.052	0.111
TAYLOR	0.024	0.063	0.119	0.015	0.063	0.117	0.015	0.055	0.107
TSE	0.064	0.133	0.189	0.061	0.138	0.211	0.065	0.141	0.195
STCC	0.012	0.052	0.103	0.012	0.058	0.102	0.011	0.048	0.099

Note: The number of replications is $S = 2000$. For $NEURAL$ the number of hidden units is $p = 20$ and the number of lags of residuals is $q = 3$, and for $TAYLOR$ the order of the Taylor expansion is $k = 3$ and the number of lags of residuals is $q = 2$.

Table 9: Small sample powers of constant conditional correlation tests under nonnormality

T	**1000**			**1500**			**2500**		
	1%	**5%**	**10%**	**1%**	**5%**	**10%**	**1%**	**5%**	**10%**
DCC – t – GARCH									
NEURAL	0.283	0.450	0.536	0.374	0.526	0.619	0.490	0.643	0.727
TAYLOR	0.511	0.668	0.739	0.636	0.772	0.834	0.788	0.885	0.922
TSE	0.113	0.223	0.329	0.183	0.338	0.436	0.245	0.432	0.530
STCC	0.246	0.380	0.461	0.301	0.427	0.493	0.339	0.455	0.529
DCC – Beta – t – EGARCH									
NEURAL	0.601	0.736	0.806	0.68	0.807	0.857	0.779	0.862	0.901
TAYLOR	0.792	0.886	0.928	0.878	0.942	0.961	0.965	0.986	0.994
TSE	0.107	0.256	0.376	0.188	0.349	0.471	0.284	0.463	0.567
STCC	0.459	0.548	0.595	0.594	0.663	0.704	0.591	0.669	0.720
DCC – Beta – t – EGARCH*									
NEURAL	0.676	0.801	0.851	0.762	0.855	0.897	0.854	0.917	0.943
TAYLOR	0.856	0.931	0.959	0.917	0.967	0.978	0.982	0.994	0.997
TSE	0.156	0.251	0.342	0.193	0.292	0.386	0.224	0.338	0.430
STCC	0.569	0.643	0.690	0.658	0.729	0.759	0.753	0.813	0.842

Note: The number of replications is $S = 2000$. For $NEURAL$ the number of hidden units is $p = 20$ and the number of lags of residuals is $q = 3$, and for $TAYLOR$ the order of the Taylor expansion is $k = 3$ and the number of lags of residuals is $q = 2$.

Table 10: Summary statistics of the daily asset returns

	Mean	Med.	Max.	Min.	S.D.	Skew.	Kurt.
JPM	0.01	-0.02	22.39	-41.13	2.75	0.91	8.08
KO	0.00	0.02	13.00	-69.56	1.46	0.01	8.12
XOM	0.02	0.06	15.86	-67.07	1.67	0.18	8.25

Table 11: Constant conditional correlation tests of the daily asset returns

	KO-JPM		JPM-XOM		XOM-KO		JPM-KO-XOM	
	Stat.	**p-v**	**Stat.**	**p-v**	**Stat.**	**p-v**	**Stat.**	**p-v**
$NEURAL$	21.8	0.001	11.8	0.067	10.9	0.015	53.8	6e-05
$TAYLOR$	39.7	0.005	73.7	8e-08	63.8	3e-06	126	2e-06
TSE	3.48	0.062	4.36	0.037	0.00	0.985	16.8	7e-04
$STCC_1$	0.91	$0.339^{(1)}$	2.92	$0.087^{(2)}$	2.44	$0.118^{(1)}$	24.4	2e-05$^{(1)}$
$STCC_2$	0.21	$0.642^{(1)}$	2.46	$0.117^{(1)}$	6.53	$0.011^{(1)}$	9.08	$0.028^{(3)}$

Note: Stat. and p-v represent respectively the value of the test statistic and the associated p-value. For $NEURAL$ the number of hidden units is $p = 20$ and the number of lags of residuals is $q = 3$, and for $TAYLOR$ the order of the Taylor expansion is $k = 3$ and the number of lags of residuals is $q = 2$. The transition variables used in $STCC$ are the lagged absolute daily returns over seven days for $STCC_1$, and the contemporaneous squared daily returns for $STCC_2$. We choose in each case the transition variable maximizing the value of the statistic (or minimizing the corresponding p-value): (1), (2) or (3) for the first, the second or the third one (see Silvennoinen and Teräsvirta 2015 for the definition of the transition variable).